VISIT…

LANZAROTE Caliente.com

全国中小学教师继续教育
教材

汉字学概要

教育部师范教育司组织评审

主编 王 宁



北京师范大学出版社

责任编辑　唐正才
封面设计　孙　琳



ISBN 7-303-05846-X/H·254
定价：8.00元

全国中小学教师继续教育教材

汉字学概要

教育部师范教育司组织评审

主　编　王　宁
编著者　（按姓氏笔画排列）
王　宁　王立军　齐元涛
宋继华　陈淑梅

北京师范大学出版社

图书在版编目(CIP)数据

汉字学概要/王宁主编．—北京：北京师范大学出版社，2001.6
全国中小学教师继续教育教材
ISBN 7-303-05846-X

Ⅰ．汉… Ⅱ．王… Ⅲ．汉字-文字学-概论-中小学-师资培训-教材 Ⅳ．H12

中国版本图书馆CIP数据核字(2001)第043168号

北京师范大学出版社出版发行
(北京新街口外大街19号 邮政编码：100875)
出版人：常汝吉
北京师范大学印刷厂印刷 全国新华书店经销
开本：850mm×1 168mm 1/32 印张：6.375 字数：158千字
2001年7月第1版 2001年7月第1次印刷
印数：1～5 000 定价：8.00元

前　言

全面推进素质教育，是当前我国现代化建设的一项紧迫任务，是我国教育事业的一场深刻变革，是教育思想和人才培养模式的重大进步。实施“中小学教师继续教育工程”，提高教师素质，是全面推进素质教育的根本保证。

开展中小学教师继续教育，课程教材建设是关键。当务之急是设计一系列适合中小学各学科教师继续教育急需的示范性课程，编写一批继续教育教材。在教材编写方面，我司采取了以下几种做法：

1. 组织专家对全国各省（区、市）推荐的中小学教师继续教育教材进行评审，筛选出了200余种可供教师学习使用的优秀教材和学习参考书。

2. 组织专门的编写队伍，编写了61种教材，包括中小学思想政治、教育法规、教育理论、教育技术等公共必修课教材；中小学语文、数学，中

学英语、物理、化学、生物，小学社会、自然等学科专业课教材。上述教材，已经在1999年底以《全国中小学教师继续教育1999年推荐用书目录》（教师司［1999］60号）的形式向全国推荐。

3. 向全国40余家出版社进行招标，组织有关专家对出版社投标的教材编写大纲进行认真的评审和筛选，初步确定了200余种中小学教师继续教育教材。这批教材，目前正在编写过程中，将于2001年上半年陆续出版。我们将陆续向全国教师进修院校、教师培训基地和中小学教师推荐，供开设中小学教师继续教育相关课程时选用。

在选择、设计和编写中小学教师继续教育教材过程中，我们遵循了以下原则：

1. 从教师可持续发展和终身学习的战略高度，在课程体系中，加强了反映现代教育思想、现代科学技术发展和应用的课程。

2. 将教育理论和教师教育实践经验密切结合，用现代教育理论和方法、优秀课堂教学范例，从理论和实践两个方面，总结教学经验，帮助教师提高实施素质教育的能力和水平。

3. 强调教材内容的科学性、先进性、针对性和实效性，并兼顾几方面的高度统一。从教师的实际需要出发，提高培训质量。

4. 注意反映基础教育课程改革的新思想和新要求，以使教师尽快适应改革的需要。

中小学教师继续教育教材建设是一项系统工程，尚处在起步阶段，缺乏足够的经验，肯定存在许多问题。各地在使用教材的过程中，有什么问题和建议，请及时告诉我们，以便改进工作，不断加强和完善中小学教师继续教育教材体系建设。

教育部师范教育司

2000 年 11 月 1 日

序　言

近年来，汉字问题引起国内外的广泛关注，这是信息时代必然产生的现象，也是文化、科技热的必然反映，应当说是一件大好事。但是，也正因为如此，社会上关于汉字的说法五花八门，违背常识的各种所谓“理论”不胫而走，造成了思想的混乱，也干扰了一些应用部门的工作。汉字水平的提高——包括对汉字科学认识的提高和正确使用汉字能力的提高，涉及到全民素质的提高。因此，在基础教育领域，识字教学科学化的问题已经成为全社会讨论的热点。古文字考据、汉字信息处理、语言文字规范与管理等领域，急需一批高级人才，如果没有汉字基础知识普遍的提高，这批高级人才的选拔和深造，数量和质量都很难达到要求。由于汉字学在信息时代应用这样广泛，近年来，它的发展十分迅速，时代需要它走出书斋，首先进入基础教育领域，然后遍及应用领域。正因为如此，在中小学教师继续教育领域，开一

门汉字学专题的课程，已经是迫在眉睫。这部教材，就是应这一需要编写的。

我们在编写这部教材时，力图达到以下要求：

第一，尽量采用最新的研究成果，但要突出基础知识。近年来，汉字学的研究在许多方面均有新的进展，比之20世纪80年代，不论在理论上还是在实践上都有不少突破。作为新世纪初编写的汉字学教材，在观点和基本理论上，虽然不能把那些前沿知识全部引进，但对已经定论或接近共识的内容，尽量不要再沿袭旧说。80年代以来，关于汉字学通论的优秀著作已经不少，例如，裘锡圭先生的《文字学概要》、王凤阳先生的《汉字学》等，都是影响很大又论证深入的著作，我们这部教材在总结汉字学基础知识时，除了阐发自己的理论观点外，也尽量学习和接受他们的一些观点，将这些论著中适合继续教育的部分综合进去。

第二，基本理论要解俗、说透，但应结合应用，有针对性，不事空谈。1998年，我们通过问卷调查了解汉字应用领域对汉字基础知识掌握的情况，发现相当大量的汉字使用者和教学者，在一些常识问题上，对各种谬误甚至荒唐的说法，完全没有或比较缺乏辨识能力，有时还不自觉地将

这些说法引进课堂。这个调查结果提醒我们，对有关汉字的基本理论，一定要解俗、说透；同时，我们要时时关注应用领域的实际问题，针对一些普遍的认识误区多加阐发，讲深讲透，少作不切实际的空谈。

第三，不求面面俱到，普及度比较高的知识，这里不多占篇幅。这部教材是为中小学教师“汉字学专题”课作的，应当适应继续教育和专题课这两个特点，比较普及的一些知识，这里就不必多说了。例如，传统的“六书”理论，前四书的知识已经非常普及，转注、假借在汉字学基础理论里实在没有讨论的必要，所以这部教材我们着重介绍了适用于各历史层面汉字的“结构—功能”分析法，转到“六书”时，也就一笔带过了。

第四，沟通古今，重视基础理论的通用性。汉字在漫长的历史进程中不间断地发展演变至今，汉字学的基础理论应当能够解释历代的汉字；而且，不了解汉字的历史，很难正确了解汉字的今天；所以，这部教材是兼通古今汉字的，既有汉字纵向的发展的描述与演变规律的概括，又有提供给汉字共时分析的操作原则和方法。

以上四点，是我们编写的主导思想，但是限于我们的水平和对实践领域了解的深入程度，加

之教材的编写需要配合课程的开设，时间很紧迫，这些想法不一定能体现得很好。不过，只要主导思想不出问题，其他方面的不足是可以逐步补足的。我们在这里首先阐明主导思想，就是为了就教于专家和读者，以利今后改进和完善。

这部教材是分头执笔的，具体分工如下：

王　宁：第一章，第二章第一、四节，
　　　　第六章第三节

王立军：第二章第二、三节

齐元涛：第三章、第四章

陈淑梅：第五章，第六章第一、二节

宋继华：第七章

全书由王宁统稿。

希望这本教材既能提供给老师教学，又能提供给已经初步掌握汉字学基础知识的学习者阅读。

王　宁

2001.3.3

目　　录

第一章　汉字与汉字学 ………………………………（1）
第一节　汉字的性质及其在世界文字中的地位 ………………………………（1）
第二节　汉字与汉语的关系 ………………………………（5）
第三节　汉字学 ………………………………（8）
第二章　汉字的起源与发展 ………………………………（12）
第一节　汉字的起源 ………………………………（12）
一、建立汉字起源假说的原则和依据 ………（12）
二、世界文字产生时的一般状况 ………………（14）
三、有关汉字起源的记载和传说 ………………（17）
四、有关汉字起源的考古资料 ………………（22）
五、关于汉字起源时代的推测 ………………（26）
第二节　古文字阶段的汉字 ………………………（27）
一、殷商甲骨文 ………………………………（28）
二、西周金文 ………………………………（35）
三、战国文字 ………………………………（40）
四、小篆 ………………………………（44）
第三节　今文字阶段的汉字 ………………………（49）
一、隶书 ………………………………（49）

二、楷书 …………………………………………………… (53)
第四节 汉字发展的一般规律 ……………………………… (55)
第三章 汉字的构形与构意 ………………………………… (59)
第一节 汉字因义构形的特点 ……………………………… (59)
第二节 汉字的书写单位和构形单位 …………………… (63)
一、汉字的书写单位 ……………………………………… (63)
二、汉字的构形单位 ……………………………………… (67)
第三节 汉字的结构 ……………………………………… (73)
一、平面结构和层次结构 ………………………………… (73)
二、汉字的构形模式 ……………………………………… (76)
三、传统的汉字构形分析方法——"六书" …… (79)
第四节 汉字构形与构意的分析 ………………………… (82)
一、汉字形义统一的条件 ………………………………… (82)
二、分析构意的基本方法 ………………………………… (85)
第四章 汉字的构形关系和使用关系 …………………… (91)
第一节 汉字的构形关系 ………………………………… (91)
一、异写字 ………………………………………………… (91)
二、异构字 ………………………………………………… (94)
三、传承字 ………………………………………………… (96)
四、同形字 ………………………………………………… (98)
第二节 汉字的使用关系 ………………………………… (101)
一、同音借用字 …………………………………………… (101)
二、职能分化字 …………………………………………… (104)
三、同源通用字 …………………………………………… (107)

第五章 汉字的规范化 …………………………… (109)
第一节 汉字的正字法 …………………………… (109)
一、历史上的正字法 …………………………… (110)
二、新中国成立后的正字工作 ………………… (112)
三、其他正字工作 ……………………………… (120)
第二节 国家的正字标准 ………………………… (121)
一、《简化字总表》 ……………………………… (121)
二、《第一批异体字整理表》 …………………… (123)
三、《印刷通用汉字字形表》 …………………… (124)
四、《现代汉语通用字表》 ……………………… (125)
第三节 书写规范 ………………………………… (125)
一、字形规范 …………………………………… (126)
二、如何避免写错字 …………………………… (129)
第六章 汉字与汉字教学 ……………………… (134)
第一节 我国汉字教学的历史传统 ……………… (134)
一、汉字教学的起源 …………………………… (134)
二、科学的童蒙识字课本 ……………………… (136)
三、对识字方法的研究和探讨 ………………… (139)
第二节 我国当前的汉字教学 …………………… (139)
一、基础教育中的教学字表研究 ……………… (140)
二、汉字扫盲教学及其教学用字研究 ………… (144)
三、对外汉字教学及其标准 …………………… (147)
第三节 汉字科学对汉字教学的指导 …………… (147)
一、根据汉字的属性来确定识字教学初期
积累的字表 …………………………………… (148)

二、依据汉字构形规律，科学地讲解字理 …… (149)
三、根据不同教学阶段和汉字不同
的属性，选择不同的教学策略 ………… (153)
第七章　汉字与信息处理 ………………………… (154)
第一节　汉字信息处理的过程 ………………… (155)
第二节　汉字编码及其评价 …………………… (158)
第三节　汉字信息处理用汉字部件规范 ……… (162)
一、汉字科学对现代汉字部件
拆分的指导 ……………………………… (163)
二、部件规范的实施步骤与标准 …………… (168)
第四节　汉字信息处理的未来发展 …………… (177)

附表一　汉字基础部件表 ………………………… (179)
附表二　笔画排序汉字基础部件检索表 ………… (186)

第一章
汉字与汉字学

第一节 汉字的性质及其在世界文字中的地位

汉字是记录汉语的书写符号。如果从殷商甲骨文开始算，汉字已经有了三千四百多年的历史，它书写了中华民族的历史，载负了光辉灿烂的中华文化；它具有超越方言分歧的能量，长期承担着数亿人用书面语交流思想的任务；它生发出篆刻、书法等世界第一流的艺术；在当代，它又以多种方式解决了现代化信息处理问题而进入计算机，迎接了高科技的挑战。汉字是中国文化的瑰宝。

世界上的文字分为两个大类：表意文字与表音文字。这种分类是从文字形体直接显示的信息是语义还是语音，也就是从文字构形的依据来确定的。例如，英语 book 直接拼出了意义为“书”的这个词的声音而成为这个词的载体。汉语“册”则用皮纬穿竹简的形态表达了它所记录的书册一词的意义而成为这个词的载体。讨论文字的性质类别，不是只考虑个别或一部分字符的构造，而要看整体构形系统总的发展趋势。汉字在发展中始终坚持保存自己的意义信息，并从个体表意，发展到系统表意。不论是表意文字还是表音文字，它们的职能都是记录语言，是语言的载体。音与义既然是密不可分的语言的两大要素，这两种文字当然同时记

录了语言的音与义，不可能只记录音或只记录义。在这一点上，表意文字和表音文字并无区别。为了不把文字记录语言的职能和它构形的依据混淆，更准确的称谓应当说，英文是拼音文字，汉字是构意文字。

有些理论认为，世界文字发展要经历表形（象形）、表意、表音三个阶段，从这个理论出发，它们认为表意文字落后于拼音文字。这个说法首先是不符合世界文字发展事实的。世界文字大都起源于图画文字，但是并不一定都经历三个阶段，表音和表意是图画文字发展的两大趋势。世界上许多古老的文字，例如非洲的古埃及文字，西亚的美索不达米亚的楔形文字等，都经历过由图画文字向表意文字发展的过程，但是，这些表意文字很快就失去了使用价值，变得不可释读了，它们虽然经历了三个阶段，但是就发展趋势而言，向表音发展是其趋势。汉字也起源于图画文字，但是，延续图画象形文字的发展趋势是表意，在数千年的历史发展中，汉字顽强地维护着自己的表意文字特点，一方面又不断地为了适应被它记录的汉语而进行了内部调整，成为世界上最古老的、具有严密系统的表意文字。像汉字一样的表意文字还有不少，它们虽然没有汉字这么长的发展历史，但也绝不是“三阶段论”所能概括的。我们主张世界文字发展“两种趋势论”，不同意“三阶段论”以及由此引发的“汉字落后论”。

就汉字记录汉语的单位而言，汉字属于音节—语素文字。也就是说，在一般情况下，一个汉字记录一个音节，而一个音节往往又代表一个语素。例如，“人”这个字记录了 rén 这个音节，而这个音节代表“人”（一种高级动物）这个语素。比较而言，英语的情况则迥然不同，英文字母代表的是音位。例如 thing（东西）这个词，th、i、ng 分别代表 [θ]、[i] 和 [ŋ] 这三个音位，thing 是三个音位拼合的词。说汉字属于“音节—语素”文字，也是从它的构形单位与语言单位的对应来说的。汉字构成以后，每个汉字

并不一定都记录一个语素，例如，在“布加勒斯特”这个译音词中，五个汉字才记录一个语素。

就形体的特点而言，汉字是在一个两维度的平面上来构形的，所以人们常称它为“方块汉字”，这个特点也是汉字的表意性带来的。早期汉字是由象形文字过渡来的，所以独体字大多呈两维平面，合体字大多采用形合的方式组成的。这种组合需要采用上下左右的相对位置来反映事物的关系。且看下面的几个例子：

例①：天


例②：目


例③：祖


例④：浴


例⑤：洗


例⑥：沬


例⑦：祝


例⑧：客


例⑨：葬


例①～③在甲骨文里都属于独体象形字，它们是直绘其形，自然置于两维平面中，例④～⑨都属于合体象形字，“洗”、“浴”、“沬”三个字，为了描写出人在器皿中洗脚、洗澡、洗脸的情境，在组字时，器皿一定放在下面，被洗的手脚应放在器皿中，而用手洗脸的人则应放在器皿边上，脸又要放在器皿近上方。“祝”绘一个人举手向木表祈祷，“客”绘人用脚走入室内用嘴说话，“葬”绘人在棺椁之中，为了体现事物的本来情境，组构的字保持了事物相互的实际位置，构形必然是一个两维的平面，而不能是线性的。所以，在古文字时代，方块汉字的格局就已经形成了，当汉字发展到义合组字和义音组字后，由于整体构形已经经过了一

番规整，当然也就要保持上下、左右的两维方形。因此可以说，方块汉字正是汉字的表意性必然带来的特点。

这个两维度的空间，为汉字构件的结合提供了很多区别的因素，除了不同的构件可以组合成不同的汉字外，相同的构件也可以构成不同的汉字：

构件的相对位置可以区别不同的汉字。例如：

“杳”与“杲”不同，“加”与“另”不同，“本”与“末”不同，“棗（枣）”与“棘”不同等等。

构件的置向也可以区别不同的汉字。例如：

“比”、“从”相异等等。

构件的数量也可以区别不同的汉字。例如：

“日”、“昌”、“晶”不同等等。

这些区别变化，都是在两维的空间内进行的。这与拼音文字由字母线性排列而结合也是不一样的。

汉字的这些特点，都是在表意文字的基础上形成的。它所以能形成和保持这样的特点，是因为这些特点可以适合它所记录的汉语。汉语是词根语，又是分析语，词形上极少有需要用音素来记录的附加成分和语法形式。汉语的语素又以单音节为主，差不多每一个单音节语素都具有意义，可以为构字提供依据。因此，汉字长期保持这些特点，不是偶然的，是汉字自然发展趋势使然。

汉字从历时的演变出发，可以分成两大阶段。这两个阶段以秦代的小篆作为分界：自甲骨文到秦代的小篆，通称古汉字；自秦汉隶书以后，通称今汉字。古今汉字的重要区别是书写单位笔画的形成。在古文字阶段汉字的书写单位是各种各样的线条，这些线条是“随体诘诎”而形成的，由这些线条构成的汉字，带有较明显的图形性。而隶、楷阶段的今文字的书写单位，则是各种类型的笔画。这些笔画经过自然发展和人为规范，逐渐变得样式固定、数量有限、写法规范，由这些笔画书写出的汉字，原始的

图形性已淡化了。

当代正在使用的汉字，称作现代汉字，现代汉字在形制上也属于今汉字。在中小学教学中，现代汉字专指已经规范了的当代汉字而言。

作为记录汉语的汉字，用自己的形体，承负着所记录的单音节语素的音和义，因而成为形音义的结合体。形、音、义被称为汉字的三要素。

第二节　汉字与汉语的关系

前面已经说过，汉字是记录汉语的符号系统，二者的相互依存关系是十分明显的：

一方面，由于汉字是汉语的载体，所以，汉字的音和义，是汉语语素音与义的反映。在汉字三要素中，有两个要素实际上是属于汉语的，唯有形，才属于汉字本体。汉语在没有汉字的历史阶段，仍然可以存在，用口语的方式，实现它作为社会交际工具的作用；但是，如果没有汉语，汉字就没有存在的必要，也没有产生的条件。从这个意义上说，汉语是第一性的，汉字是第二性的。

正因为汉字是第二性的，所以，正如前面说过的，汉字的特点，是适应汉语的特点而形成与保存的，同时，汉字的发展演变，在很大程度上要受到汉语的制约与推动。例如，中国周秦时代是汉语单音语词大量孳生的时代，词的派生大大推动了汉字的孳乳；也正是在这个时代，汉字的形声化趋势急剧发展，形声字大量产生。魏晋以后，由于汉语的造词方式逐渐变为以合成为主，汉字的增长速度也就逐渐缓慢了。

面另一方面，汉字一旦产生，它对汉语使用和发展的影响也

是不容忽视的。汉字不但帮助汉语超越空间与时间的限制，扩大了交际职能，而且它具有超方言的特点，因而有利于民族共同语的形成。同时，由于有了汉字所记录的书面语，更促进了汉语的严密化。汉字作为汉语的视觉符号，对分辨汉语中的同音词、吸收古语词、方言词和外来词，都起了非常重要的作用。

尽管汉字与汉语相互依存，但是从本质上，汉语并不等同于汉字，它们是互有差异的两种符号体系。

汉字记录汉语的单音节语素，在古代汉语里，汉字与单音词在很大程度上是对当的，这就容易造成一种错觉，认为汉字和汉语可以等同。这个认识是不全面的。将汉字与汉语混为一谈，不论在理论上，还是在实践上，都会产生失误。

汉字和汉语的区别可以从以下几个方面来看：

第一，汉字虽然由于记录汉语而从汉语的语素那里移植了音和义，但是它还有属于自身的形式——字形。因此，汉字除受汉语的制约与推动外，同时又有属于它自己的、不受语言制约的发展变化规律和使用规律。汉字在发展中要逐渐形成自己的构形系统，构形系统的总体对汉字个体是有制约作用的。很多与构形系统不相切合的异体字被自然而然地淘汰，很多新产生的形声字对义符和声符的选择，都是汉字构形系统的内部规律使然。汉字的构形系统与汉语的音义系统不是同一个系统，这一点对教学有很大的启示：依靠汉字构形系统集中识字，往往难以设计思想内容切合小学生的课文；而利用言语作品分散识字时，又往往难以完全切合汉字的难易程度。这种现象，就是汉字构形系统与汉语词汇、语音系统不一致带来的。

第二，文字和语言不是同一时期产生的，在讨论它们的历史发展时，不能把二者混为一谈。例如，在汉字中，构形比较单一、理据比较清晰的独体象形字，比由它构成的合体字特别是形声字产生要早；但是，这并不意味着独体字所记录的词都一定早于形

声字所记录的词。在汉字中，“马”是独体字，它的产生早于形声字“妈”，但是就词而言，“妈妈”这个词，并不比“马”产生晚。从总体看，汉字史与汉语史有密切关系，但并不完全等同。

第三，在汉语的书面语言作品中，字与语素或词的对当关系是不平衡的、不整齐的。在一般情况下，说一个字记录的是一个语素或一个单音词，不会出太大的问题。但是在分析书面文献时，对文字与语言、字与语素或词的不平衡状态，不整齐对应的状态，就要特别引起注意：

由于异体字的存在，不同的作品甚至同一部作品中常常出现异字同词的现象。例如同样是“痛苦”的“痛”，汉代的碑文上时常写成“慟”。同样是“危险”、“高危”的“危”，魏晋作品中时常写作“峗”。

由于同音借用字的存在，同一语素写两个字和不同语素写一个字的现象都不乏见。例如：

“从容”、“容易”、“容纳”的“容”和“容貌”的“容”在意义上毫不相干，却写同一个字。

“雅座”、“雅量”、“雅语”的“雅”写同一个字，但“雅座”的“雅”是“迓”的借字，当“迎接”讲；“雅量”的“雅”古代常写作“疋”，当“酒杯”讲；“雅语”的“雅”一般认为是“夏”的借字，“雅语”即“夏语”。

“马寅初”的“马”是在中国可以追寻历史渊源的姓氏，“马克思”的“马”只记录一个德语音节的声音。

在虚词里，字与词不一致的情况更为普遍。例如，韩愈的《马说》最后两句话：“其真无马邪？其真不知马也。”第一个“其”与反问情态副词“岂”一样，可译作“难道”；第二个“其”是商榷情态副词，可译作“恐怕”、“该不是”。

在共时平面上是如此，在历时情况下更为复杂，形体上具有传承关系的字，在意义上并不完全相当。例如，甲骨文的“天”、

“大”同用，周代金文已经分开，先秦典籍“房舍”与“舍弃”的“舍”不分，隋唐时后者已经写作“捨”；现代简化字又都写作“舍”了。

由于以上原因，在教学和研究中分清字与词就十分重要。不但字词考据要注意这个问题，统计时也要注意，当我们要在一篇文章或一部文献里统计共有多少词或语素时，绝不可以它的字数来替代。上述《马说》的两句话中，共有不重复的词 9 个，但仅有不重复的字 8 个。

这些事实都说明，在理论上分清汉字与汉语本质上的不同，在实践中注意字与语素、词的差异，是非常重要的。

第三节 汉 字 学

前面说过，表意文字和拼音文字是世界文字中并存的、代表着两种发展趋势的文字系统，它们各有其特点，又各有其发展规律。汉字在几千年的发展历史中，一直坚持着表意的特点，不停顿地被使用至今，成为世界上惟一的一种有着日渐严密构形系统的表意文字，是表意文字的代表。正因为如此，汉字的发展演变中，有拼音文字不具有的现象和规律，也有其他发展时间较短的表意文字未曾出现过的现象和规律。所以，研究汉字的构形特点和使用规律，不仅是中国文字学的课题，而且是世界文字学的课题。

汉字学是以汉字为研究对象建立起的一门学科，研究汉字的学科从汉代的“小学”算起，已经有了一千多年的历史。但是汉字学成为独立的学科时间并不太长。传统“小学”虽然早在隋唐时代从名目上已经分为训诂、体势、音韵三类，以后又形成了文字、音韵、训诂学科三分的局面，但是就学科内容的划分来说，音

韵学确实成为独立的学科，文字和训诂却迟迟未能分割清晰。这是因为，汉字本体的研究必须以形为中心，而且必须在个体考证的基础上探讨其总体规律，历时数千年的传统文字学在研究上以形附属于义、着重个体而忽略总体的习惯，便无形之中成为这种本体研究的障碍。加之历代的文字研究都依赖字书，而古代字书都不区分字形的历史层面，提供不出一批经过整理的系统字料，总结汉字构形的整体系统便更加难以起步。同时，古代的学科重视综合和应用，这种特色更加使文字学与训诂学的界限难以釐清，作为汉字本体的研究没有和对它功能的研究分开而有其独立的体系。20世纪初发展起来的古文字学，不但使汉字构形的规律逐步得到了多方的印证，而且使文字断代的观念得到了强化，这样才有了汉字学独立的可能。

汉字学的基础理论是从传统“小学”、古文字学和近年兴起的现代汉字学中总结出来的最基本的规律。它应当具有这样的特点：第一，它是从汉字的实际材料中总结出来的，而不是照搬某些现成的结论；第二，它是对具有继承性的历代汉字进行断代的测查而形成的，因而有着切合汉字发展每个历史阶段并能说明汉字发展演变总规律的普遍性，而不是只对一部分或某一阶段的汉字适合；第三，它不但能解释发生在个体汉字上的诸多现象，而且能把汉字看成是记录汉语的符号系统加以描写并对现象作出解释；第四，它应有彻底的理论性，不是就事论事，因而对汉字的整理和规范等应用领域可以提供一些可操作的科学规则。按照这几个条件来建设汉字学的基础理论，是大家努力的方向。

这门学科发展到今天，实际上形成了以下四个方面的分支：

第一，汉字构形学。探讨汉字的形体依一定的理据构成和演变的规律，包括个体字符的构成方式和汉字构形的总体系统中所包含的规律。就汉字的发展历史来说，不同历史阶段的汉字构形具有各自的特色，而汉字构形学要能涵盖各阶段汉字构形的诸多

现象，为研究各阶段汉字提供基础理论和基本方法。汉字构形学是其他几个分支的枢纽和基础。这是因为，通常所说的汉字三要素形、音、义中，音和义都是汉字作为汉语的载体由汉语那儿承袭来的，只有字形是汉字的本体。不论研究汉字的字源、字用、风格和它所携带的文化信息，都必须先把汉字的构形规律搞清楚。

汉字形义学与汉字构形学是从不同角度提出来的。这种研究从理论上说，是要抓住汉字因语素的意义而构形的特点，总结出汉字形义统一的规律，在此基础上，探讨如何通过对汉字形体的分析达到确定它所记录的词的词义这一目的。从实践说，是要借助字形的分析来探讨古代文献的词义，为古书阅读和古籍整理提供语言释读的依据。

汉字构形学与汉字形义学是一项工作的两个方面。前者借助于意义，探讨的中心是形体，所以属于汉字学范畴；后者借助于字形，探讨的中心是意义，所以属于训诂学或文献词义学范畴。

第二，汉字字源学。尽量找出汉字的最早字形，寻找每个字构字初期的造字意图，也就是探讨汉字的形源，也叫字源，这是汉字字源学的任务。字源学是研究探讨形源的规律和汉字最初构形方式的学科。

汉字字用学与汉字字源学是从不同角度提出来的。个体字符造出后，并不是永远用来记录原初造字时所依据的那个词或词素，它的记录职能时有变化。字用学就是研究在具体的言语作品里汉字字符记录词和词素时职能的分化和转移的。

汉字字源学探讨原初字形，属于汉字学范畴，字用学探讨汉字记录汉语的实际职能，属于训诂学或文献词义学范畴。

第三，汉字字体学。汉字字体指不同时代、不同用途（鼎彝、碑版、书册、信札等）、不同书写工具（笔、刀等）、不同书写方法（笔写、刀刻、范铸等）、不同地区所形成的汉字书写的大类别和总风格。研究汉字字体风格特征和演变规律，是汉字字体学的

任务。

第四，汉字文化学。这种研究有两方面的目的：一方面是宏观的，即把汉字看成一种文化事象，然后把它的整体，放在人类文化的大背景、巨系统下，来观察它与其它文化事象的关系，这是宏观汉字文化学；另一方面则是微观的，即要研究汉字个体字符构形和总体构形系统所携带的文化信息，对这些文化信息进行分析，加以揭示，这是微观汉字文化学。总之，汉字文化学是在作为文化事象的汉字与其他文化事象的互证关系中建立起来的，在解释汉字这个中心任务上，汉字文化学是对汉字构形学的补充。

汉字学这四个分支的内容是互相联系、密不可分的。

第二章
汉字的起源与发展

第一节　汉字的起源

汉字的起源要说明从没有汉字如何到有了汉字，因此，它的资料必然是未经汉字记录过的，它的结论不能一一验证，只能是建立某种假说。科学的假说也是有充分依据的，其中的推论也要符合逻辑。这方面，现代考古学家、历史学家、文字学家都做出了一些成绩，虽然有些说法还存在分歧，但在大的框架上，已经有了不少共识。这里择其要介绍。

一、建立汉字起源假说的原则和依据

历史发展的总规律是越远古越缓慢，人类经过漫长的蒙昧野蛮时期，逐渐进入了文明时代。现今史学界把“文明”一词用来指一个社会氏族制度解体而进入有了国家组织的阶级社会，而文明的重要标志之一便是文字的产生和应用。因此，我们从一般逻辑判断，一种文字总是在蒙昧通向文明的转折点上诞生，这应当是合乎历史发展规律的。

世界上文明的发生都在大河流域，因为，水源和动植物给人

类的生存提供了可能的生态条件。中国的文明在黄河流域诞生，它是独立发展的。汉字——中华文明的重要标志之一，也是独立产生的，它和埃及圣书字、古代苏美尔文字、原始埃兰文字和克里特文字等，同是世界上最古老的文字。汉字和这些古老的文字一样，都经过由图画文字到表意文字的阶段；而它与其他古老文字不同的是，那些古文字在演变中有的停止使用而丧失了生命力，有的变成了拼音文字，有的甚至不可识读，被外来文字取代，唯有汉字，没有停顿地被使用至今，成为世界上惟一的一种有着日渐严密体系的表意文字。

既然汉字发展的历史没有中断，那么，我们沿着它踏出的足迹向上追溯其起源，应当是可能的。但是，现存最早的汉字甲骨文产生在殷商时代，距今大约有三千四百年的历史。它已是一种有着比较严密体系的文字。小屯的殷墟文化，仅从它相当发达的冶炼技术和铜器、陶器、玉器上的艺术纹饰看，便可以确定它已是高度发达的文明。所以有人说，如果把殷墟文化看作中国文明的诞生，那就未免有点像传说中的老子，生下来就有了白胡子。近年来，在西安市郊区长安县花楼子客省庄二期文化遗址的发掘中，发现了一批刻在骨笄、兽牙和兽骨上的刻符，据说已经清理出单体文字十多个。这一文化遗址属龙山文化晚期。有人把这些刻符也称作甲骨文，从而把甲骨文产生的时代推前了一千多年。但是，从汉字学的角度看，殷墟甲骨文是一个成熟的文字体制的专称，在没有证实长安刻符与殷墟文字的历史源渊关系时，我们不能把凡是刻在甲、骨上的符号都称作“甲骨文”，而把它和殷墟文字挂在一起。

我们完全可以推测，甲骨文产生的时代，恐怕不会处在中华民族由蒙昧走向文明的路口。汉字在进入甲骨文之前，必定还有一段相当长的历程。但是，沿着甲骨文再向前追溯，可以证实汉字历史面貌的文物与文献还不够系统和充分，因此，对汉字的起源，只能以不连贯的历史证据和依照发展逻辑的推测为基础，建

立科学的假说。

建立关于汉字起源的假说，首先应当弄清，原始汉字处于什么状态就算已经产生了。考古学家们认为，只要发现了一种有标示作用的符号，就可以确定为汉字起源的初阶。文字学家们认为，文字是一种记录语言的符号，原始的标记符号必须脱离了任意绘形、任意理解的阶段，产生了一批具有约定的意义、可以记录语言中的词——也就是有了固定读音——的单字，并且可以开始进行字料的积累的时候，才能算已经产生了。要达到这个状态，必须具备两方面的条件：一方面是社会发展的条件，社会生产力发展到一定的程度，生产组织进一步严密，居住点与政治中心确立，继承法定，才使口语的局限变得日益突出，超越时间与空间限制的文字才不得不发展起来；另一方面是文字发展本身的条件，文字正式产生以前，必须有一批可以书写、可以传递的图画或符号，可以用来演变为字符，有了这些可以称作“前文字现象”的材料为基础，才有创造文字的可能。我们综合考古学家和文字学家的看法，在讨论汉字起源问题时，取一个时间上线与下线的幅度。这个幅度或许会有一两千年，但这对于人类的发展来说，仍然是十分短暂的。

汉字究竟何时起源？起源时的情况如何？是根据以下四种材料来进行推断的：

（1）根据世界文字产生、发展的一般规律来推断；

（2）参考对历史有折射反映的神话传说；

（3）根据仅有的与汉字起源有关的书面文献记载；

（4）根据已经发掘出的与汉字起源有关的考古资料。

二、世界文字产生时的一般状况

人类生产发展到一定的程度，异地的交往扩大，异时的记忆

量增多，口头语言的局限急待克服，这时，首先可以用来传递信息的是具有记号性质的实物。结绳和契刻这两种记事方法几乎遍及各大陆。根据记载，古埃及、古波斯、古代日本都曾有过结绳之事。人类学家和民俗学家考察，近代美洲、非洲、澳洲的土人，我国的藏族、高山族、独龙族、哈尼族……也都有用结绳记事的风俗。秘鲁的土人用数条不同颜色的绳，平列地系在一条主要的绳子上，根据所打的结环在哪条绳上、什么位置和结、环的数目，来记载不同性别、不同年龄的人的人口数。这表明，结绳确实是历史的遗存。至于刻契，主要用于战争和交换，在超越空间限制传递信息上，功能大于结绳，但记事的功能更弱，历史的普遍性也较差。不过，以齿记数的作法，对文字的产生也会有启发，作为前文字的一种现象，也是可信的。我国西南某些少数民族（如基诺族），在四十多年之前还一直使用刻木记事的方法。但这些方式只能是对自身记忆或双方默契的提示，对毫无关系的两个人，不加有声语言的解释，便不具有交流思想与交际的性质。有些实物经约定后可以作为固定意义的信号。例如，中国古代的虎符，必须契合才有传递命令的作用。《孟子》记载，依照周礼，召唤守苑囿的虞人要用皮冠，如用旌召唤，虞人可以不去。这些，都是原始人实物传递信息的方式在军事、政治上的残留。这种方法，也只能用于特定的场合，传达极为单一的意义。

一种可以用于较大范围又能传递较多信息的手段，必须使接受信息的人与发出信息的人想到的是同样的事物和内容。在人类文明的起点上，唯有写实性图画可以起到这种作用。因为，写实性图画是现实生活的再现，它具有独立的说明性，而且，不同的图画所具有的区别性，可以使不同的信息得以区分。所以图画用于传递信息，是文字的源头。而用图画来传递信息，开始时有些近似绘画艺术的作用，靠的是它的直观性和形象性，绘形越细致，给予对方的信息量越大也越准确，越能沟通信息的发出者与接受

者。下面的四幅连续图画被称作图画文字信。

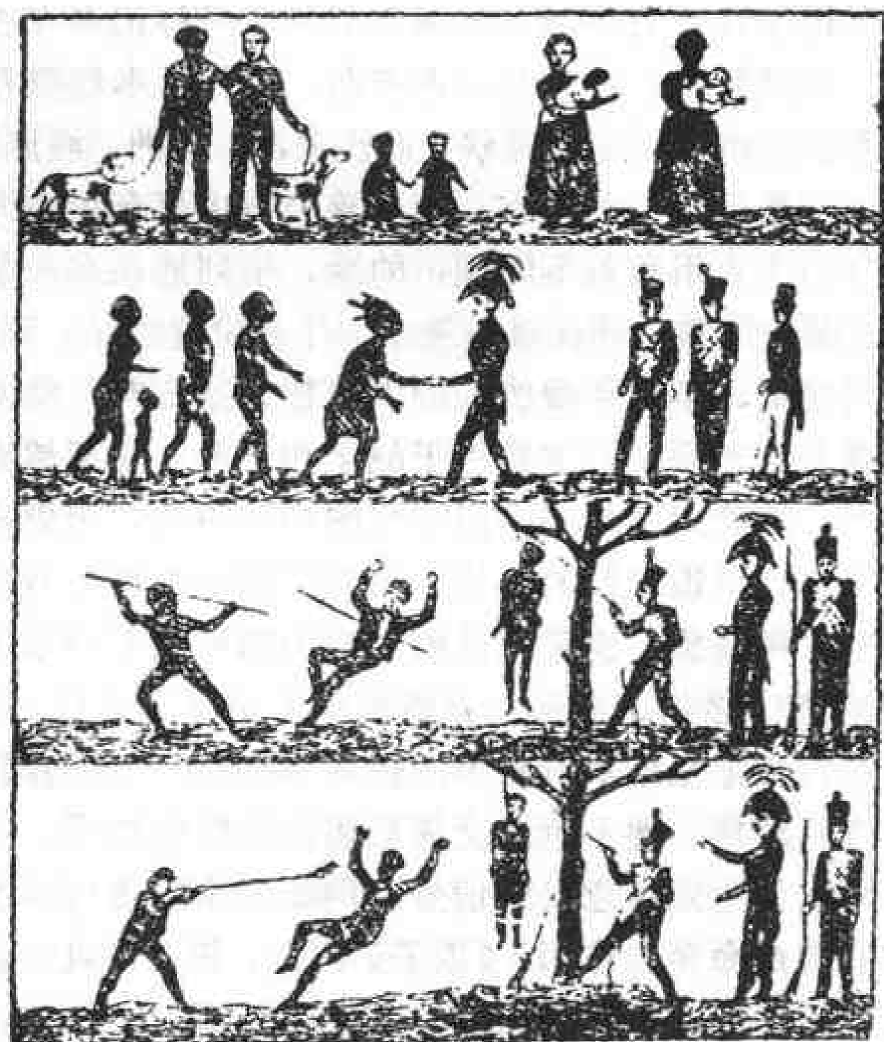


图 2-1　塔斯马尼亚总督给土著人的图画文字信，建议和解，同时指出破坏和约者会遭遇的后果。①

可以看出，它是用以传递信息的图画，并没有获得文字的资格。因为它是凭着完全写实的图形，使收信者从直观上理解写信者的意思的。经过信息传递的多次重复，使某一绘形与某一意义建立了固定的联系，形意关系带有了约定性，这才有了图画文字

① ［苏］B·A·伊斯特林著．文字的产生和发展．北京：北京大学出版社，1987．79

的性质。图画文字具有的形意关系，不是与语言对应的，而是超语言的。也就是说，这些符号可识而不可读。人们可以用不同的语言单位，选择不同的词句去指称它的意义。例如，商代的铜器上，常常有一些图像，是作为征伐的标志的。只要看到这种众人聚在旗下的形象，就可知这一铜器与征伐有关。但它并没有与哪句话或哪个词固定地建立起关系。因此，它还属于文字画或图画文字的性质。

图画文字发展为早期象形文字，最主要的变化，在于它直接成为语言的符号，变得可识可读。另外，它的总体的直观性经过分析，变为一形即一词。这时，完全写实的绘形就演变为象征的或局部替代的，只要所绘与所指保持理据，能使人识别就可以了。图 2-2 是北美印第安人用来表达抽象概念的象征性图像。它已摆脱了写实，析成单图，只是尚未与词对应，有意而无音。因此，它可以看作由图画文字向早期象形文字过渡的一种形态。

在图画文字发展为象形文字的过程中，原始的花纹图案也是远古时代约定符号的来源之一。在花纹图案中，某些表义成分常常转化为图画字，然后又转化为表词字。在东欧草原地带发掘的公元前 20 世纪末至前 10 世纪初期木椁文化的陶器，上面有许多象征性的花纹图案，例如圆圈象征太阳，波形线象征水，十字形象征四面八方……它们与以后产生的文字常常有渊源关系。

三、有关汉字起源的记载和传说

迄今为止，我们所能看到的有关汉字起源的文献记载，最早来自周秦的典籍，而且大都是传说，有的还带有神话色彩。但是，神话与传说有它产生的历史背景和根源，往往是某些历史现实的折射反映。关于汉字起源的历史旧说，并不都是无稽之谈，其中

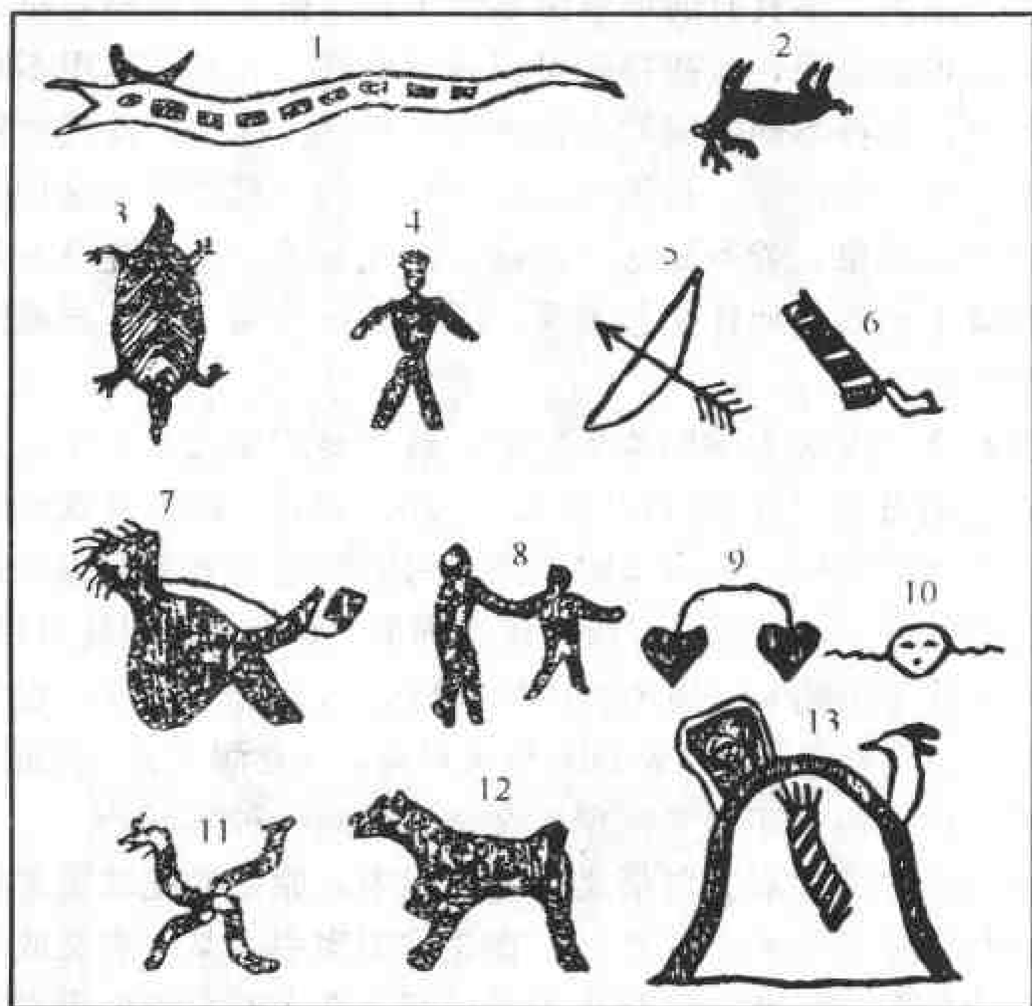


图 2-2　北美印第安人图画文字中用来表达抽象概念的象征性图像①

1. 生命（神话中有角的蛇）　2. 死亡（头朝下的动物或人）　3. 幸福、成功（龟）　4. 灵巧（双翼代替两手的人）　5. 战争（带箭的弓）　6. 和平（插着羽毛的和解烟斗）　7. 讲和（人吸和解烟斗）　8. 友谊（联在一起的手）　9. 爱情（联在一起的心）　10. 注意、听（两耳边有波形线的头）　11. 危险（两条蛇）　12. 保卫（狗）　13. 祈祷（举手朝向天及诸神）

① ［苏］B·A·伊斯特林著. 文字的产生和发展. 北京：北京大学出版社，1987. 71

也有一部分是信传。尽管这些传说并非汉字起源历史面貌的详尽写实，但是，它们对推断汉字起源前后的状况和推断汉字起源的大体时间，都是很有价值的。

有关文字起源的传说，大体可分成两类：一类是关于前文字时期的传说，另一类是关于创造文字本身的传说。从这些传说中，我们可以窥见原始汉字发生的因由，以及由原始汉字向成熟的文字体系过渡时的一些历史状况。在这些传说中，结绳说与仓颉造字说最值得我们注意。

汉字产生以前，中国曾有用实物记事的阶段。结绳和刻契是其中的重要手段。结绳记事的说法首见《易经·系辞》："上古结绳而治，后世圣人易之以书契。百官以治，万民以察，盖取诸夬。"《庄子·胠箧篇》也说："昔者容成氏、大庭氏、伯皇氏、中央氏、栗陆氏、骊畜氏、轩辕氏、赫胥氏、尊卢氏、祝融氏、伏牺氏、神农氏，当是时也，民结绳而用之。"根据这两个说法，上古有很长一段时间都用结绳记事，神农氏是用结绳的最后时代。至于结绳这种记事方法如何施行，《周易正义》引《虞郑九家易》说："古者无文字，其有约誓之事，事大大结其绳，事小小结其绳，结之多少，随物众寡；各执以相考，亦足以相治也。"人们把结绳与文字联系在一起，是因为人类创造结绳记事的方法与发明文字的想法是很一致的。一件事情要想保留在人的脑子里，只有在记忆所能达到的时间和准确度之内，才是可能的。但记忆的延续时间和可负荷的容量都是有限的，只有用外部的标志来提示它。人们在相互约定某一事情后，也需要有一种客观凭据以便长期遵行。前者是记录某种思想内容，后者是记录某种交往内容，使这些内容超越时间的限制，这正是激发人类发明文字的动因。也就是说，到了结绳时代，文字产生的主观要求已经具备了。结绳的可区别性很低，只能用结大结小来标记大事小事，像秘鲁土人用不同颜色的绳并在一起，再加上颜色与结绳位置的区别，最多也只能传递

十几种至几十种信息，所以，它的记录功能是很弱的。但是，它既有治事的作用，后来又用于约誓之事，能够“各执以相考”，因而也必然具有一种约定的内容；原始社会的人群活动范围还不很大，对记事符号的交际功能要求不高，突破语言的时间限制比突破空间限制更迫切一些，结绳作为一种视觉的记事符号，在记事的数量和明确性上虽然极为寥寥，但它是一种成功的尝试。从结绳到文字，虽然发展了几千年，但在性质上，距离已不很遥远。结绳说既有后代民俗作为确证，又可从历史发展的逻辑上分析出它存在的合理性，因此，时间错乱的“八卦”说，虚玄的“河图洛书”说，是不可能与结绳说同日而语的。

在有关创造汉字的传说中，仓颉造字说也是一种有价值的传说。这种传说最早出现在战国时代的文献里。《吕氏春秋·君守》说：“奚仲作车，仓颉作书，后稷作稼，皋陶作刑，昆吾作陶，夏鲧作城，此六人者，所作当矣。”《荀子》、《韩非子》也有关于仓颉造字之说。到了秦汉时代，仓颉造字说流传更广，影响更深。《淮南子·本经训》有“昔者仓颉作书而天雨粟，鬼夜哭”的传说。李斯统一文字时所用的课本，第一句就是“仓颉作书”，所以称作《仓颉篇》。把前人传说吸收后加以整理，正式写人早期汉字史的是东汉的许慎。他在《说文解字·叙》里说：“及神农氏结绳为治而统其事，庶业其繁，饰伪萌生。黄帝之史仓颉，见鸟兽蹄迒之迹，知分理之可相别异也，初造书契。”又说：“仓颉之初作书，盖依类象形。”《文心雕龙·练字》沿袭许慎的说法，才有了“文象立而结绳移，鸟迹明而书契作”的名句。

对于仓颉造字说，过去的历史学家们着眼于考证仓颉是否实有其人，如果有，大约在哪个年代。由于确凿史料的缺乏，很难得出结论。而在我们看来，仓颉造字说的可取之处，主要在于它说出了汉字起源的一些道理，许多是很有参考价值的。

首先，这种传说把结绳与仓颉造字衔接起来，认为在“庶业

其繁”以后，结绳无法适应更多、更快地记录、传递信息的需要，人们必须探索新的方式，创造更多的相互区别的符号，来记录更多的信息。在“兽蹄鸟迹之道，交于中国”的时代，人们从鸟兽蹄迒之迹得到了“依类象形”、“分理别异”的启示，逐渐创造了文字。这个说法是可信的。

其次，传说仓颉是黄帝的史官，也是有道理的。文字产生在国家形成过程中，首先是政事往来的需要。所以，汉字形成过程中起主要作用的应是与文字有密切关系的巫史。史与文字的关系，也可从汉字构形中得到证实：

甲骨文“史”作 ，后来分化为“史”、“吏”、“事”三个字。《大盂鼎》的“卸事”（御事），卜辞称“卸史”，是殷周治事之官。卜辞中“贞史”是“问事”，“史贞”是“事问”，正见“史”、“事”是一字。关于“史”字的构形，有许多不同的说法，有一种说法是：从“中”，“中”是簿书、典册。“又”是手，以手执中正是史官的形象。《礼记·礼器》：“因名山升中于天。”是讲祭天时烧柴，置玉帛于其上，同时也连文册一起烧掉。“中”是文册。《周礼·秋官·小司寇》：“岁终，则会群士计狱、弊、讼登中于天。”“登中于天”就是写在天府的登记册上，“中”即“册”。《周礼·春官·天府》：“乡州及都鄙之治中，受而藏之。”郑众注：“治中谓其治职簿书之要。”所以，“史”是书写、收藏簿书的官，他们是直接并大量使用文字的人。

仓颉是史官，因集中使用原始文字，得以对群众自发产生的字符加以规整。《荀子·解蔽》说：“故好书者众矣，而仓颉独传者，壹也。”对“壹”字有不同的理解，有人解释作“专一”，把《荀子》这句话理解成：“仓颉只是众多好书者中由于用心专一而最有成就的一个。”又有人以“壹”为“同一”、“统一”，认为《荀子》这句话应当理解为仓颉整理过文字。其实，《荀子·解蔽》所说的“壹”指正道，也就是正确的规律。“壹”与“两”相

对，不受邪说的蒙蔽叫作“壹”。荀子认为，后稷之于稼，夔之于乐，舜之于义，和仓颉之于书一样，都是因为专门从事某方面的工作，从面掌握了正确的规律，才能独传。仓颉是一个因为集中使用文字而摸着了它的规律从而整理了文字的专家。在汉字从原始文字过渡到较为规范的文字的过程中，他起了独特的作用。可以推断，这样的一个人，在汉字起源阶段的晚期，一定会存在的。

“仓颉”出现的时代,应在原始汉字有了一定数量的积累阶段，也就是中华民族由蒙昧走向文明的初期。说他在神农氏之后的黄帝时代，是因为黄帝代表中华民族共同的祖先，黄帝时代是中华文明的发源时代，至于具体的时间，是不足为据的。

四、有关汉字起源的考古资料

传说毕竟不是确凿的史料，汉字起源的确切时代和起源的状况，曾经那样模糊。近六十年来，现代考古的谜一个个从地下揭晓。关于汉字的起源，也有了更多的考古资料显示了蛛丝马迹。

殷墟甲骨文产生在商代后期，约为公元前14～前11世纪。这时的汉字已经是可以较好地记录汉语的文字系统。这种汉字的实际面貌表明，它不但经过相当长时期的积累，而且经过某种程度上的整理规范。那么，在此之前的汉字是什么样子的呢？

从商代后期再往前，推到商代前期。考古学家认为，郑州商代遗址的发掘中找到的被称作二里岗文化的商代文化，与殷墟文化是一脉相承的。它的时代大约在公元前1750—前1350年范围内，早于安阳小屯殷商文化三百多年。在翻转的地面上，曾发现三件有字的骨，其中两件各有一字，另一件有十个字。考古学家认为是为练习刻字而刻的，它的形制与殷墟甲骨文差不多，可以看出，当时已经在使用甲骨文。二里岗文化的陶器和陶片上也刻划符号，都是简单的刻文，刻在陶大口尊上。河北藁城县台西商代

遗址里也发现了一些刻在陶器上的符号，其中一部分比商代后期略早一些。江西清江吴城遗址出土的陶器刻符，也早于殷墟四百年左右。这二处的陶纹，有的与汉字很相像。如图 2-3。

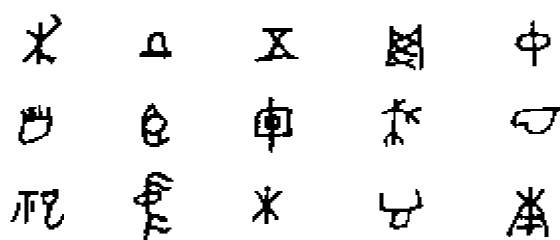


图 2-3 考古发现的陶纹

虽然这些符号都不连贯成句，相互似无关系，但它们对汉字的起源，都能起到一些参证的作用。

近年来，对夏文化的探讨日渐热烈，一般认为夏代已经进入阶级社会，起码应该有一些原始文字，但是在夏代遗址的发掘中，却没有找到可以确证为文字的资料。河南偃师县发掘的二里头文化，根据考古层位，是压在河南龙山文化层之上，又被二里岗文化所压住的。它的时代约在公元前 19～前 15 世纪，属夏文化，至少它的晚期已经进入了文明阶段。在发掘的陶片上，有的刻划记号用意不清。诸如：



图 2-4 陶片上的刻画记号

这些几何式的直纹，应当都与文字有关。

尽管商代早期和夏代文化中似乎没有更多的关于汉字的考古资料，但不能说明那时没有文字。骨刻、金铸、陶绘的器物上得以保存文字，是因为它们的质地不易腐烂。可以想见，在汉字有所积累的时期，它不可能不用来进行日常书写。那种临时性的交往或保留时间不必过长的书写物，是可能划在易于腐烂的兽皮、木片、树皮、树叶、草叶等东西上的。而这些书写物在数千年之后无法再现，就像我们已经无法看到原始社会的木船和草鞋一样，应

当是十分正常的。

而且，隔开这一段，在更早一些的新石器时代，我们还可以窥到一些文字起源的信息。

山东的大汶口文化，发现刻有符号的陶尊和残器片十六件，大都出自墓葬，上面有十八个刻符。这些刻符的图像和甲骨文与早期金文的象形文字非常接近。试比较：

大汶口陶符	甲骨文	早期金文
	甲文“旦”	金文“旦”
	甲文“山”	金文“山”
	甲文“斤”	甲文“新”从此
	甲文“戌”	
	甲文“丁”	金文“丁”

图 2-5　大汶口陶符与甲金文的比较

虽然这些符号绝大多数刻在陶尊外壁口下颈部，而且一个陶尊一般只刻一个符号，只有一件在颈部两侧刻两个，另一件在颈部、底部各刻一个，相互并不连贯，因此，意义无法探求。但说它们是原始文字，已有相当的理由。

属于仰韶文化早期的半坡遗址，发掘出精美的彩陶，它的时间大约在公元前 5000—前 4500 年。彩陶在中原地区到龙山文化时期便衰退了，但却在黄河上游的甘肃、青海地区得到了发展。马家窑文化与半山马厂文化约在公元前 3000—前 2000 年，都有图案华丽的彩陶。彩陶上的花纹一般认为是装饰性的美术图案。近两年来，有人对庙底沟彩陶纹中的鱼形纹发生了兴趣，认为它是带有表意性的由鱼文化遗存的符号。甲骨、金文中有些字很像是由鱼纹演变来的：

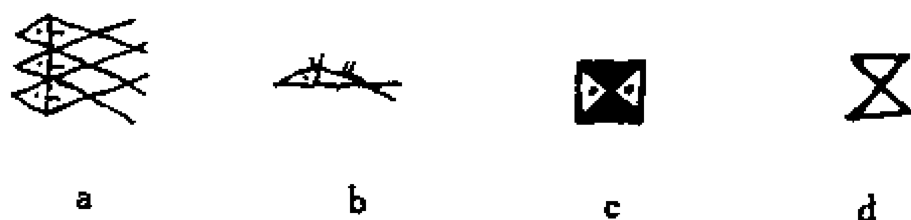


图 2-6　彩陶中的鱼纹分解之一

图 2-6 中，a 是彩陶中套画的鱼图案，劈成单项成 b，两个相对的鱼头拼成 c，表示相交，最后演变成 d，即甲骨文的“五”字。“五”、“午”同源，确有相交之义。



图 2-7　彩陶中的鱼纹分解之二

图 2-7 中，a 是彩陶中套画的鱼头图案，组合为 b，是四个鱼头相对，简作 c、d。春秋金文“四”作，正是从 c 变来；战国简“四”作，正是从 d 变来。

一说，甲骨文、金文的“明”字，也是鱼头纹演变来的：

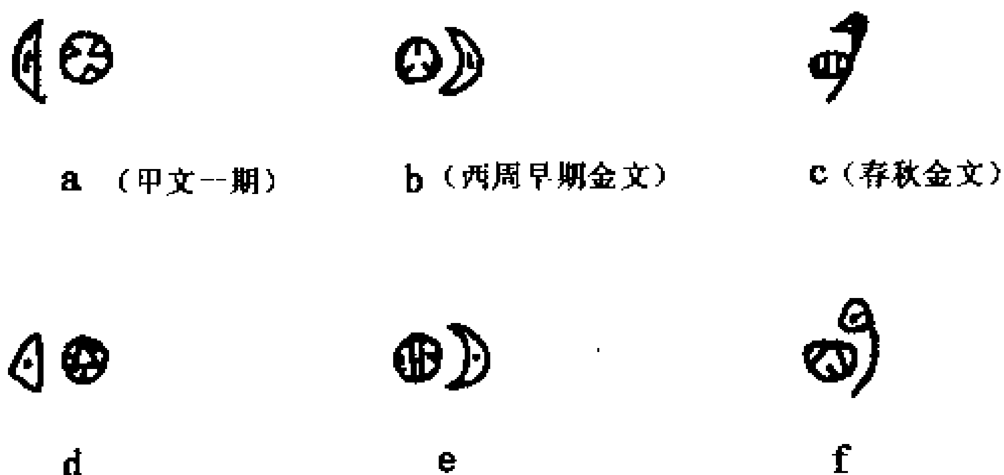


图 2-8　彩陶中的鱼纹分解之三

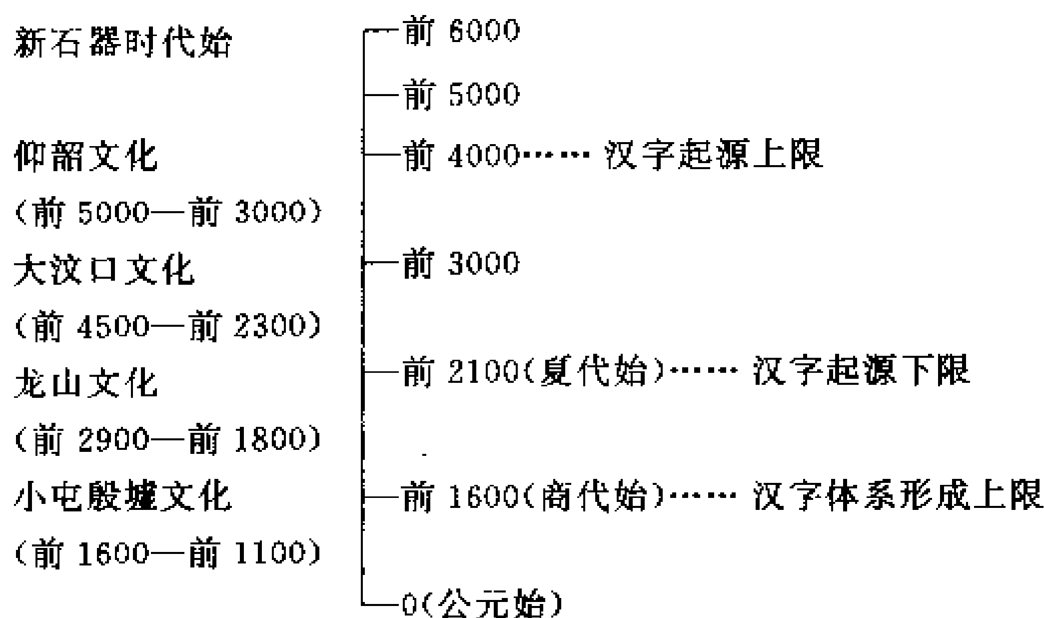
图 2-8 中，a、b、c 是甲骨文、金文，d、e、f 是鱼头纹，它们是何等相似。“明”正是用鱼目来表示明亮。所以“盟誓”的“盟”字甲骨文作，金文作，都只画一边，不画另一边，而小篆“明”作，正是变来的。旧石器时代人们以渔猎为生，而鱼对黄河、长江流域以及滨海原始居民的生活，更有特殊作用。新石器早期的鱼形纹与后来文字的关系，确有不容忽视的一面。

五、关于汉字起源时代的推测

由图画传递信息到产生图画文字，再由图画文字和花纹图案的表义成分转变为约定符号以至记词符号，这需要一个相当长的过程。这就是文字起源的过程。欧洲和美洲的一些古老文字的这个过程，大约始于新石器时代而止于有史时期的开始。从所获得的资料看，汉字的产生就其上下限而言，也正是在此时期内。

到目前为止，我们在已出土的文物中，没有看到作为汉字前身的用图画传递的信息和图画文字，因此，汉字起源的上限难以从实际上确定下来。现在能够提出的根据，最远的只有公元前4000 年左右属于仰韶文化的彩陶纹，以及时间与此相近的属于大汶口文化的陶器刻符。它们的时代，都在新石器时期的中期，可以暂时把这一时期作为汉字起源的上限。从理论上说，汉字起源的上限也许比这还要早些，那要等发现新的考古证明时再向上推移了。至于汉字起源的下限，我们可以从小屯殷墟甲骨文往上推测。当图画文字与陶器花纹中的表意图案演变为记词字符，汉字由零散的、个别的字符逐渐积累，达到一定的数量后，再通过人为规范，成为一种文字体系。这一过程，如果跟记词字符产生的那段过程相比，应当是短一些。这不仅是因为，生产和政治的发展都越来越迅速，更因为起源阶段具有多元性（由于部落不同，可能有多种原始文字存在）、自发性、群众性的特点，而初期发展阶

段却是在少数集中使用汉字的人手中自觉进行的。小屯殷墟甲骨文已是能够完整记录汉语的文字体系，这个体系形成的开端应当在夏商之际。《尚书·多士》记载西周初年周公的话说："惟殷先人有册有典，殷革夏命。"这就是说，商人在灭夏时，已经有了记事典册。从古史文献也可以看出，夏代是中国第一个有完整世系流传下来的朝代。那么，汉字字符开始积累的年代，似可估计为夏初，也就是公元前 2100 年左右。



从上表可看出，原始汉字在新石器时代中期开始产生，到它发展成初步的文字体系，大约经历了近 2600 年。

第二节　古文字阶段的汉字

汉字从形成体系发展到现在，已经有了三四千年的历史。在这漫长的历史长河中，随着社会的进步和汉语的发展，作为记录汉语的汉字符号系统，从字符的构形到书写体势，都有过几次重大的变化。按照以往的习惯，汉字的形体演变可划分为两大阶段，

即：古文字阶段和今文字阶段。秦代小篆属于古文字的下限，古文字主要包括殷商甲骨文、两周金文、战国文字以及秦代小篆，有人认为秦隶（古隶）也属于古文字；汉代隶书属于今文字的上限，今文字主要包括汉隶至今的文字。

一、殷商甲骨文

甲骨文是刻在龟甲、兽骨上的文字的通称。商代和周代都有甲骨文，但现已发现的甲骨文大部分属于殷商时期，因而人们习惯上把甲骨文视为殷商文字的代表。

甲骨文最早是在位于安阳小屯的殷墟发现的，所以又称殷墟文字。起初，小屯村民耕地时发现甲骨，以为是中药的龙骨，以很低的价钱卖给了药材店。关于当时的情形，罗振常《洹洛访古游记》中有详细记载：①

此地埋藏龟骨，前三十余年已发现，不自今日始也。谓某年某姓犁田，忽有数骨片随土翻起，视之，上有刻画，且有作殷色者，不知为何物。北方土中，埋藏物多，每耕耘，或见稍奇之物，随即其处掘之，往往得铜器、古钱、古镜等，得善价。是人得骨，以为异，乃更深掘，又得多数，故取藏之，然无过问者。其极大甲骨，近代无此兽类，土人因目之为龙骨，携以视药铺。药物中固有龙骨、龙齿，今世无龙，每以古骨充之，不论人畜。且古骨研末，又愈刀创，故药铺购之，一斤才得数钱。骨之坚者，或又购以刻物。乡人农暇，随地发掘，所得甚

① 罗振常《洹洛访古游记》宣统三年二月二十三日条，上海隐庐书店 1930 版。

多，检大者售之。购者或不取刻文，则以铲除削之而售。其小块及字多不易去者，悉以填枯井。

这么多的甲骨被销毁，实在太可惜了。1899年，北京的古董收藏家王懿荣，在一个很偶然的机会发现了甲骨文。这年他得了疟疾，医生给他开的药方中有一味叫龙骨的中药。药买来后，他无意中发现"龙骨"上刻着好似篆文但又不认识的文字。他平素喜好金石学，对古文字有较深的造诣，马上意识到此"龙骨"决非一般药材。于是他开始寻踪收购，并精心研究，最后终于考定这些"龙骨"原来是殷商时代的卜骨，其上的文字是比篆籀还要早的殷商文字。

甲骨文一经发现，便得以迅速的发掘和收集。到目前为止，公

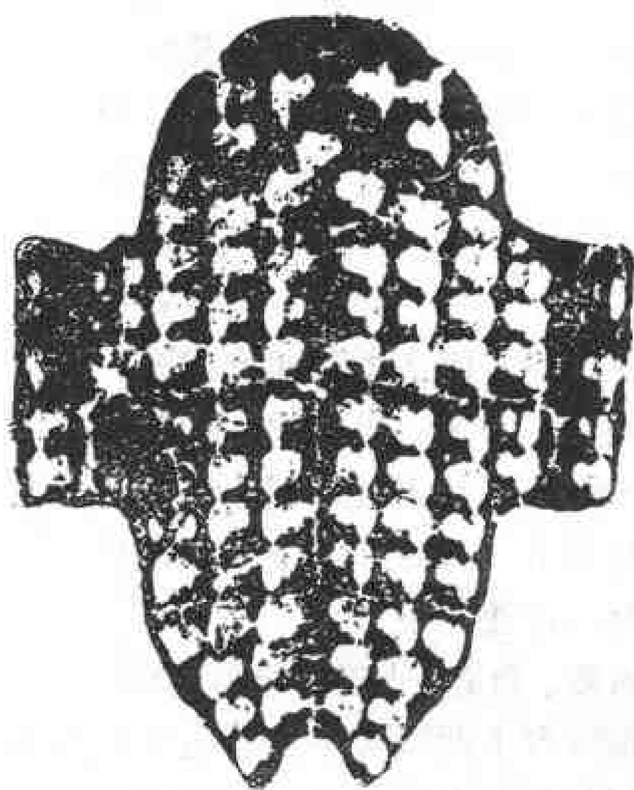


图2-9　龟腹甲背面之钻凿

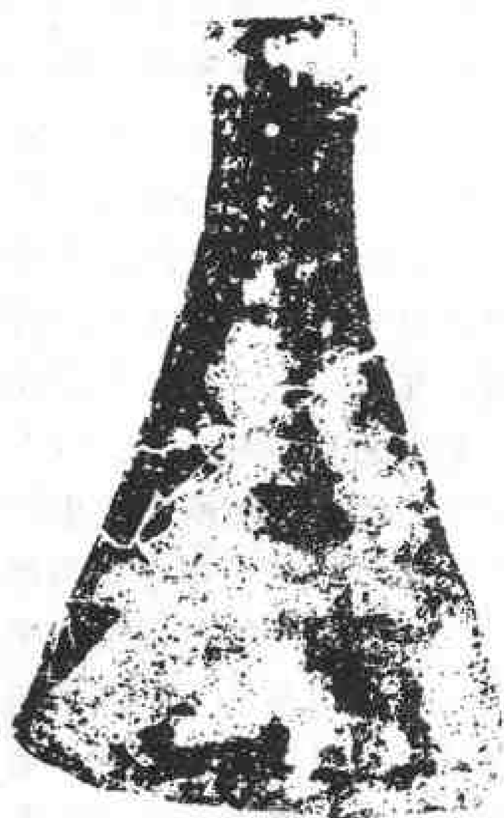


图2-10　牛右肩胛骨

私所掘总计约十五万片。目前汇集出土甲骨片最精、最全的，是中国科学院考古研究所编的《甲骨文合集》。搜集单字最完备的，是考古研究所在孙海波《甲骨文编》的基础上编写的改订本。其中正编收已识字 1 723 个，附录收未识字 2 949 个，共计收单字 4 672个。

甲骨文的内容相当丰富，但绝大部分是关于占卜的记录，所以甲骨文又称卜辞，或称殷墟卜辞。殷商时代，社会文明还比较落后，人们非常迷信，相信上帝左右一切，因而无论大事小事，如天气、收成、田猎，甚至做梦，都要卜之于天。当时的占卜材料主要是龟甲（图 2-9）、兽骨（图 2-10），这些材料需要经过加工才能使用。龟壳自腹甲与背甲之间截分为二，上下甲之间原有甲桥连接，锯开之后甲桥联属腹甲，同时锯去外缘部分使甲桥成为有规则的弧形。背甲则从中脊处一剖为二，或锯去近中脊的凹凸不平处，削去首尾两端，使成鹅卵形。卜骨要锯去骨脊及臼骨。材料准备好后，还要钻凿。钻凿是为了使甲骨某处变薄，以便在此处显兆。钻凿后的槽穴大致呈“丁”字形（图 2-9）。钻凿成这种形状，也是为了出现兆干和兆枝两条裂缝，从而构成“卜”字形；而且也可使兆纹有规律地排列，钻在凿左，兆枝在左方；钻在凿右，兆枝在右方。兆纹是通过烧灼槽穴使之爆裂而形成的，爆裂所发出的 pu 声，正是“卜”字读音的来源。当时的人们就是根据这些兆纹去判断所卜之事的吉凶的。这其实就如同小孩子玩的游戏，却被当时人视作极为神圣的事情。有时在占卜结束之后，人们还要把占卜的过程刻在甲骨上，这便是卜辞。

一则完整的卜辞通常包括前辞、命辞、占辞、验辞四部分。前辞是占卜日期和占卜者，命辞是所占卜的具体事情，占辞是根据兆纹所作出的占卜结果判断，验辞是占卜过后的应验记录。有时占卜人为了慎重，对一件事往往反复占卜，这样还要在每一兆纹之侧刻上占卜的次数，称为纪数字。如图 2-11 便是一个完整的

例子：

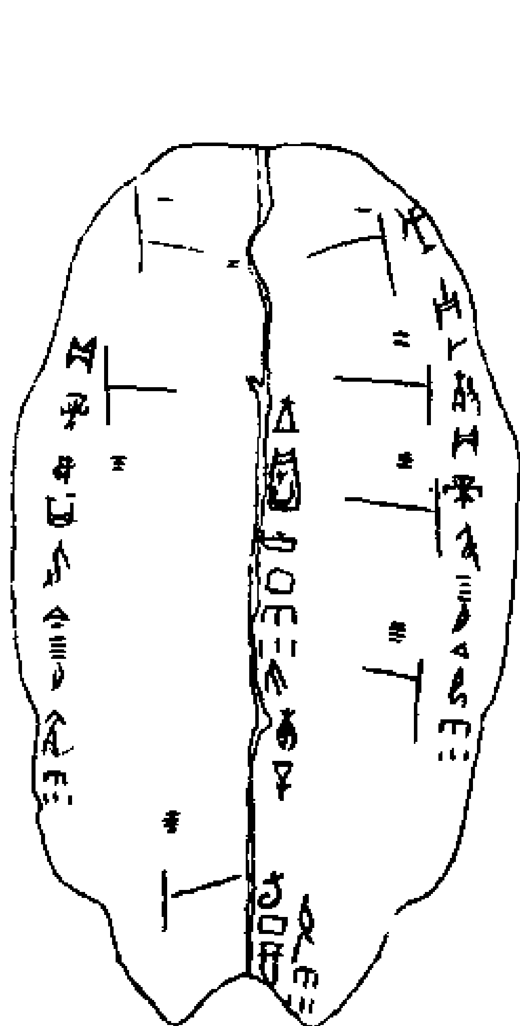


图 2-11

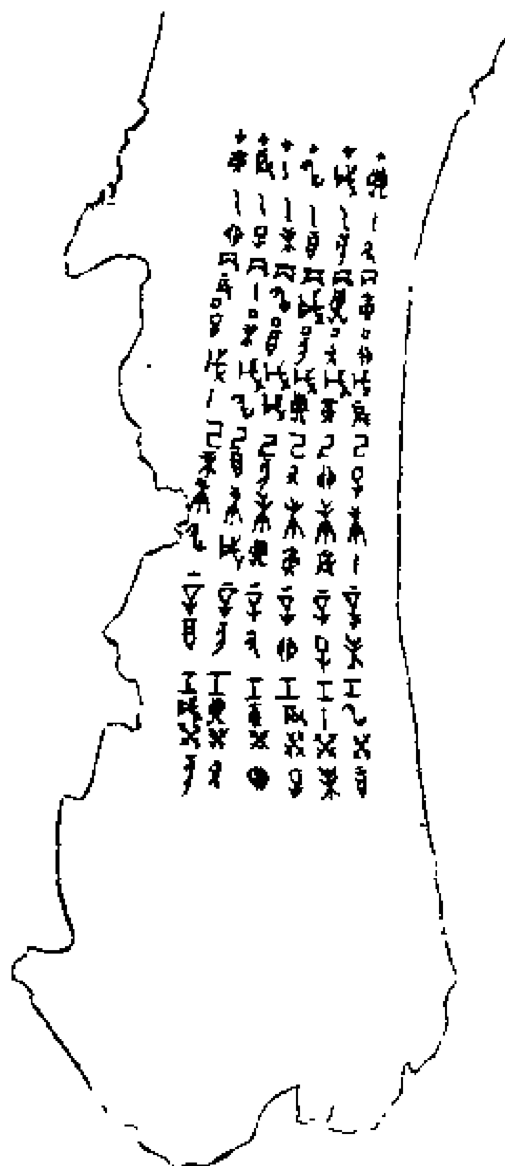


图　2-12

前　辞：戊子卜，𣪊贞。

命　辞：帝及四月令雨？

　　　　帝弗其及今四月令雨？

占　辞：王占曰：丁雨，不叀辛。

验　辞：旬丁酉，允雨。

纪数字：一 二 三 亖

也有些甲骨文不是占卜的记录。其中有些是附刻在占卜材料上、与占卜有间接关系的干支表（图 2-12），以及关于甲骨的纳贡、收藏等杂事的记录；有些则与占卜全无关系，纯属纪事，大都刻在不作占卜用的兽骨上，既不施钻凿，更无所谓灼兆。

甲骨文大部分是用刀刻的，所以甲骨文又有“契文”的别名。但也有极少数甲骨文是用毛笔写的，或墨书，或朱书。实际上，早在商代，毛笔就成了主要的书写工具。“笔”的繁体作“筆”，从竹从聿。“筆”原本作“聿”，竹字头是后来添加的。“聿”甲骨文作 ，正像以手握毛笔状，说明当时已用毛笔。《尚书·多士》说：“惟殷先人，有册有典。”“册”甲骨文作 或 等形，正象用竹片或木片编成的简册之形。可见当时就有了竹简文，只是由于竹简易于腐烂，我们今天才没能看到，因此说，甲骨并不是当时的普通书写材料，而主要是为占卜这种特殊目的而选用的特殊材料。我们今天虽然无法看到商代用毛笔书写的典籍，但从一些出土文物上，我们仍然可以看到少量的商代毛笔字（如图 2-13）。



图 2-13 商代文物上的毛笔字

与商代毛笔字相比，甲骨文笔画瘦削，多有方折，刀笔味很浓。这主要是由于书写材料的坚硬而造成的。也有个别的甲骨文笔画粗壮，雄浑庄重，与铸在青铜器上的铭文相差无几，这也许是先用毛笔写好，再按照毛笔字迹刻成的。

从构形的角度看，甲骨文已形成了相当成熟的文字体系。传统六书中的象形、指事、会意、形声等四种造字方法，在甲骨文

里都已具备。例如：

象形：（女）（豕）（门）（矢）（星）

指事：（亦）（甘）（肱）（上）（下）

会意：（伐）（及）（牧）（逐）（获）

形声：（新）（星）（鸡）（责）（狼）

这说明，当时人们对汉字基础部件的各种功能已有了初步的认识，既看到了独体字符较原始的表形功能，也懂得了将它们转化为表义或示音功能，并且能够自觉地运用各种基础部件构造新字。所有这些，都为汉字构形系统的形成和进一步完善奠定了基础。

但是，殷商甲骨文毕竟还处在汉字发展的初级阶段，在有些方面还较原始，带有明显的早期汉字的特点：

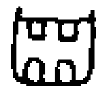
（1）此期汉字象形性很强，有些甚至带有原始图画的特征。如下列各字：



（渔）



（鸡）



（齿）

有些字究竟像何形，现在已无法得知，从而造成了大量的未识字。

（2）许多字书写置向不定。或朝左，或朝右；或向上，或向下；或正面，或侧面。如：

卜：　　至：　　龟：

（3）字形结构不固定，异构字较多。有些字采用了不同的构

件，如“牢”字；有些字构件的位置不同，如“物”字；有些字构件的多少不同，如“凤”字；有些字繁简的程度不同，如“奚”字等。

牢：　　　物：

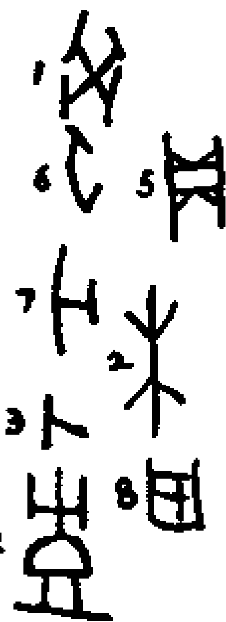
奚：　　　凤：

(4) 经常出现两个或两个以上的字合写在一起的现象。这种现象叫“合文”。如：

大乙　小臣　小祖乙　四祖丁　报丙　三万　五十　十二月

(5) 有少数字形有混同现象。如“山”和“火”都写作，“甲”和“七”都写作十等。虽然原则上讲“山”为平底、“火”为圆底，“甲”竖略长、“七”竖略短，但在实际材料中是很难分清的。

(6) 行款不固定。行款是字的刻写顺序。甲骨文的行款复杂多样，有直书，有横书，有左行，有右行。更有的简直杂乱无章，颠三倒四，不仔细推敲就根本读不通。如图 2-14，须按所标序号的顺序读才能成句，即：“癸未卜，壴贞：旬无祸?”


图　2-14

以上几个方面充分说明，殷商甲骨文虽已大致形成了相对成熟的文字体系，但它距汉字起源的时代应该不会十分遥远。

二、西周金文

金文是铸刻在青铜器上的文字的总称。钟、鼎是青铜器中乐器、礼器的代表，故金文又叫钟鼎文。钟鼎上的文字有阴文和阳文两种，阴文叫款，阳文叫识，故金文又叫钟鼎款识。青铜器常被用作祭器，祭礼古代叫吉礼，故青铜器又叫吉金，其上的文字又叫吉金文字。祭器又称彝器，故金文又有彝器铭文、彝器款识的别称。

金文的发展，与青铜器的藏礼作用有着密切的关系。所谓藏礼，就是寓礼于器，是古代宗法礼制在青铜器上的物化。青铜器早在商代以前就出现了，当时主要是一些日用器具，其上也没有铭文。商代前期，个别青铜器上开始出现族徽性的单字。商代后期，随着奴隶主宗法礼制的逐渐强化，青铜器的藏礼作用日益明显。一些经常用于祭祀宴享的青铜器被赋予了特殊的意义，具有了明贵贱、别尊卑的特殊功能。青铜器的不同组合方式代表着不同的身份和等级，如天子九鼎、诸侯七鼎、大夫五鼎、士三鼎等。青铜器成了国家的象征、家族的荣耀，“问鼎”便意味着侮蔑和挑战，“鼎迁”则意味着国家灭亡；家族成员立了战功，受了封赐，就要铸刻在青铜器上，以便昭示后人，世代为荣，故金文中常有“子子孙孙永宝用”的字样。到了西周，青铜器的这种作用达到了顶峰，其上的铭文也变得越来越长。西周末年的《毛公鼎》共铸有 497 字，是现知最长的一篇金文。西周以后，奴隶制开始向封建制过渡，“礼崩乐坏”的局面愈演愈烈，青铜器的藏礼作用逐渐衰落，到战国时期重又回复为日用器具，长篇铭文已非常罕见。

由于西周金文是在宗法礼制的推动下走向兴盛的，因而此期金文的内容多数与君臣、宗族有关。如祭祀典礼、颂扬先祖、征伐纪功、赏赐册命、训诰臣下等。图 2-15 的《师兑簋》就是典型的册命体铭文。大意是：周王在康庙里册命师兑，让他继承祖先的



图 2-15　师兑簋

官职，并赐给他一些象征身份的服饰。师兑感到非常荣耀，就将此事铸在簋上，使子孙后代永志不忘。

西周金文是直接继承殷商甲骨文而来的。与殷商甲骨文相比，西周金文有了明显的进步。具体表现在以下几个方面：

（1）由于书写材料的不同，西周金文在笔画上一改甲骨文瘦

削方折的特点，变得肥厚粗壮，圆浑丰润。字体庄重美观，大小渐趋一致。行款多为直书左行，排列也越来越整齐，有时甚至先打格、后书写。

（2）新的象形字很少出现，而形声字却大大增加。如“走”部、“言”部、“金”部、“厂”部的字在甲骨文中很少或几乎没有，而西周金文中则大量出现，并且多为形声字。形声字的大量增加，是汉字构形系统走向成熟的一个重要标志。

（3）异构字相对减少，结构渐趋定型。形旁因意义相近而混用的现象已不普遍，如“辵”部的字，在甲骨文中往往有从“彳”、从“止”、从“辵”三种写法，而西周金文则基本固定为“辵”。某些部件的置向及位置已不再随意改变，如从“彳”之字已基本将“彳”固定在左边，“及”、“方”等字的反书现象已很少出现。

（4）合文减少。西周金文虽然仍有少量合文，但数量却比甲骨文少得多，而且排列方式也不再那么随意，逆读合文消失。如“五十”写作 ，而不再象甲骨文那样写作 。

这说明，西周金文比殷商甲骨文更为成熟了。不过，西周金文中还保留着一些图画性很强的早期象形文字，有些甚至比殷商甲骨文还要古老。这种复古倾向可能是出于某种特殊的需要，如充当宗族标志的族徽文字，就需要用典雅庄重的古体以显示宗族的尊严。这些个别的复古现象，并没有阻止西周金文在总体上走向进步的步伐。

以上这些方面，是我们就整个西周时期金文的共同特点总结出来的。实际上，西周金文自武王克殷（约公元前1130—前1018年之间），至平王东迁（约公元前770年），历经几百年，其间的变化还是十分明显的。试比较图2-16《天亡簋》和图2-17《史颂簋》：《天亡簋》属西周早期铭文，笔道粗细不匀，多有填实的团块；蜿曲笔道较多，象形性较强；字体大小不一，东歪西扭，排列极不整齐。相比而言，西周晚期《史颂簋》上的铭文，笔道粗



图 2-16　天亡簋



图 2-17　史颂簋

细已相当均匀，填实性的团块也为线条所代替；字体大小基本一致，且均呈长圆型，已初具小篆形制；排列较为整齐，行列分明。

如对单个字的变化进行比较，则更能说明问题。如表一各字突出体现了填实性团块的线条化，而表二各字则显示了蜿曲笔道的平直化过程：

表一：

	王	才	辛	人	生	大	侯	父
西周早期								
西周中期								
西周晚期								

表二：

	隹	馬	無	寶	彝	周	殷	貝
西周早期								
西周中期								
西周晚期								

西周金文前后期的显著变化表明，汉字的演变并不以朝代的更替为界限，而是时刻都在进行着的。我们有时按朝代去介绍，主要是为了叙述的方便，并不是说同一朝代的文字就基本一致，不同朝代的文字就相去甚远。如，与西周晚期金文相比，西周早期金文更接近于商代晚期金文。

三、战国文字

战国文字是对战国时代周王室和各诸侯国所有品类文字的统称。它不像甲骨文、金文那样是对某一特殊品类文字的专指，也不像小篆、楷书那样是对某一定型字体的专称。我们之所以这样指称，是因为这一时期无论哪一种器物上的文字，都不能成为该期所有文字的代表。

目前所见的战国文字可分两个系统：一是见于后代文献上的战国文字，如《说文》中的古文和籀文、三体石经中的古文和郭忠恕《汗简》中的古文等；一是见于战国各类出土文物上的战国文字，如用毛笔写在竹简上的竹简文、写在丝帛上的帛书、写在玉片上的盟书，以及铸刻在金属器上的金文（包括符节文和货布文）、刻在石头上的石刻文、刻在印章上的古玺文和刻在陶器上的古陶文等。文献上的战国文字几经传抄，往往失去了原来的真面目；而文物上的战国文字则是当时的真迹，是研究战国文字最可靠的资料。

经过春秋时期的动荡，战国时期已经形成了诸侯割据的局面。正如许慎在《说文解字·叙》中所说的那样，当时“诸侯力政，不统于王，恶礼乐之害己而皆去其典籍”。对典籍的轻视，使得人们做事没了共同的依据，写字没了共同的标准，这样，便出现了“田畴异亩，车涂异轨，律令异法，衣冠异制，言语异声，文字异形”的局面。当时各国文字之间的差异，从下表中可见其一斑。

	小篆	西周正体	战国文字						
			秦	楚	齐	燕	三晋	中山	
馬	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
者	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
市	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]		
年	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

注：此表转引自詹鄞鑫《汉字略说》

战国文字的地域性差异虽然复杂，却具有一定的规律性，即：距周王朝所在地越远，变化就越大。盘庚东迁之后，秦迁都于雍（今陕西凤翔附近），这里原是西周故地。秦迁来后，基本上继承了西周文化，在文字上也与西周金文一脉相承。当时的秦国相对比较落后，政治经济上都较为保守，文字方面也没有太大的改变，除了书写风格上渐趋规整匀称之外，结构上的变化并不明显。而六国文字则变化剧烈，与西周金文差距越来越大。所以，人们往往把战国文字分作两大派系，即：秦系文字和六国文字。

秦系文字上承西周金文，下启小篆，是汉字发展主线的一个重要环节，因而它是战国文字的主流。六国文字偏离了正统文字发展的轨道，未能引导当时汉字发展的方向，因而是战国文字的支流。

目前所能见到的战国秦系文字的资料很少。其中最重要的应该是保留着战国秦文字原貌的石鼓文（图 2-18）。石鼓文是刻在十个馒头形石头上的文字。因为这些石头形状有点像鼓，故称石鼓。



图 2-18　石鼓文（局部）

其上刻有四首四言诗，格调与《诗经》相仿。原文约有六百多字，但自唐初发现以来，几经搬迁磨损，只剩下三百多字了。这些石鼓文的结构端庄严谨，大小一致，笔形布局极有法度，偏旁部首的写法和位置也都基本定型；笔道粗细均匀，已基本实现线条化（图中斑块是因残蚀所致）；字体风貌已与小篆十分近似，明显处于西周晚期金文向小篆的过渡阶段。

1979 年，在西安市郊发现了一个《秦杜虎符》（图 2-19），根据上面所说的“兵甲之符，右在君，左在杜”，可知是秦始皇后来统一天下改称皇帝以前的兵符。其上的文字线条匀称，字体长圆，已是典型的小篆了。这说明，真正的小篆战国末即已形成，并不

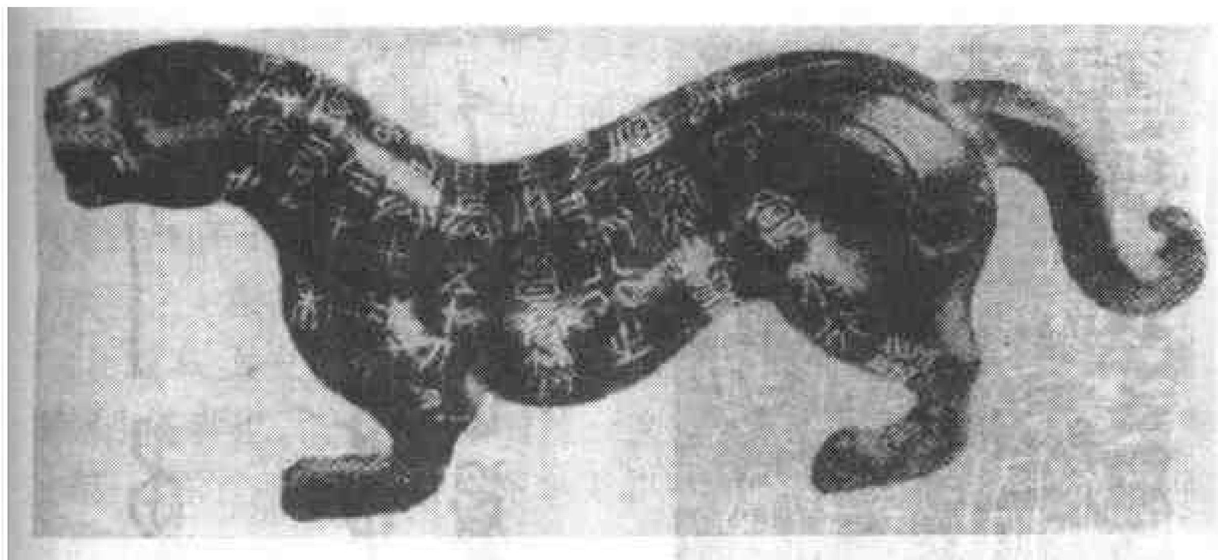


图 2-19　战国秦地虎符

是秦始皇统一后由李斯等人创制的。

与秦系文字相比，六国文字显得十分杂乱。总的来说，六国文字具有以下几个特点：

（1）地域性分歧较大。在战国以前的春秋金文中，就已经出现了某些地域性特点，但当时主要体现在书法风格上，字形结构方面还看不出明显的特色。进入战国时期，各国的字形结构产生了很大分歧，有些差别大得简直看不出是同一个字，如 40 页表中所示。

（2）同一国家和地域的异写异构现象也很普遍。如同是齐国陶文，“区”字却有多种写法：　　　　。甚至同一篇铭文中，也会出现差距很大的形体，如鄂君启节中“昭”字作　，又作　等。

（3）字形变化随意性过大，无规律可循。许多字形省变严重，已很难看出前后的传承关系，也无法从构形理据上去解释。如“马”省作　等。

（4）部分文字装饰性强。如曾侯乙编钟铭文（图 2-20），线条纤细，字体瘦长，精心美化，带有春秋时期鸟虫书（图 2-21）的遗风。



图 2-20　战国曾侯乙编钟铭文图



图 2-21　春秋兵器鸟书文字(二器)
左:"蔡公子加之用"戈　右:"玄"戈

(5) 简化倾向十分明显。这也许是因为当时的正统文字不便于快速书写，无法满足日常事务的需要，于是便对之大幅度简化。

由于六国文字对汉字乱加改造，破坏了汉字的构形理据，违背了汉字自身的发展规律，因而随着六国的相继灭亡，六国文字也逐渐退出了历史舞台。不过，六国文字虽然是汉字发展历史中的匆匆过客，但它对汉字的发展并非没有任何影响，例如它的简化倾向，对其后隶书的出现极具有启发意义。

四、小　篆

小篆是秦始皇统一中国后实行"书同文"政策时所采用的标准字体。

小篆成为正统文字，一方面得益于秦朝利用国家权力进行推

广，另一方面也是汉字自身发展规律的必然结果。

秦始皇统一中国后，为了便于对六国遗民进行统治，统一文字便成了当务之急。而要想统一文字，首先必须确定一个统一的标准。许慎《说文解字·叙》说："秦始皇帝初兼天下，丞相李斯乃奏同之，罢其不与秦文合者，斯作《仓颉篇》，中车府令赵高作《爰历篇》，太史令胡毋敬作《博学篇》，皆取史籀大篆，或颇省改，所谓小篆者也。"可见，当时统一文字的标准是秦文，凡是不与秦文相一致的，一概废除；而秦文的标准，则是由史籀大篆省改而来的小篆。

所谓"史籀大篆"，是指收录在《史籀篇》中的文字，称"大篆"，又称"籀文"。一般认为《史籀篇》是周宣王时的史官籀编写的一部童蒙识字课本。如果此说成立的话，其上的文字应为西周晚期金文。如前所述，西周晚期金文已与小篆形体非常接近，小篆在其基础上省改而来，是完全合乎情理的。不过，据近人研究，《说文》中所保存的"籀文"与西周晚期金文并不相同，而是战国时的秦系文字。对此，我们认为有两种可能：其一，《说文》中的籀文确系西周晚期金文，只是由于《说文》几经传抄，已失去了籀文的本来面目，从而与西周晚期金文有了一定的差距。如《说文》把籀文和古文都用对称与两头尖细的笔法来书写，致使二者无体势风格之别，就是证明。其二，《史籀篇》原系西周晚期金文，但流传至战国时期，秦人便改用当时的秦系文字去抄写，而《说文》所依据的本子正是秦人的抄写本，其中的"籀文"自然就与秦系文字一样了。这两种可能都仅仅是推测而已，由于资料的不足，目前还没有确切的结论。但无论是哪种可能，都不影响小篆来源于"史籀大篆"的说法，因为西周晚期金文、战国秦系文字、小篆都是一脉相承的。秦朝之所以强调小篆是从《史籀篇》而来的，主要是为了树立这种字体的威信。因为《史籀篇》是周王朝制订的，而周王朝又是各诸侯国的共同祖先，以来自周王朝的文

字作为标准去推广，很容易为六国遗民所接受，而实际上所推广的仍然是稍加省改的秦系文字。

其实，小篆并不是秦始皇统一中国后才出现的，更不是李斯等人所创制的。早在战国末年，小篆就已经基本定型了。如前面所举的秦杜虎符就充分说明了这一点，秦朝所采用的文字标准，就是这种现成的秦文字。李斯等人所做的工作，只是对小篆进行一番整理和加工，使之更为定型化、标准化，然后书写出标准字样（如《仓颉篇》等）加以推广。

小篆的流行，虽然带有一定的政治成分，但也是符合汉字发展的自身规律的。小篆在保存汉字构形理据的基础上，使汉字构形进一步简明化、系统化，而不像六国文字那样为了简化面任意破坏汉字的构形理据，因而小篆最终取代六国文字而一统天下，是历史发展的必然结果。

目前所能见到的秦朝小篆也可分文物和文献两类。文物上的秦朝小篆比较少，其中较有代表性的有《泰山刻石》（图 2-22）和《峄山刻石》（图 2-23）。其中《泰山刻石》据说是李斯手迹，可谓是标准的小篆。

文献上的秦朝小篆主要见于东汉许慎的《说文解字》。这是保存小篆最系统的一部书，共收录小篆字头 9 353 个，加上书中保存的一些籀文和六国古文，共计 10 516 字。《说文解字》是我们释读古文字的重要桥梁，也是查检小篆字形的重要工具书，在汉字学史上具有极其重要的地位。

通过对上述小篆材料的分析，我们可以将小篆的特点总结如下：

(1) 小篆已形成一个相当严密的构形系统。在甲骨文阶段，象形字占有很大比例，这些字各以其所代表的物象为构形依据，相互之间只强调区别，而缺乏有机的联系。有些象形字在添加声符而成为形声字后，仍保留着原有的物象特征，使得这类形声字独



图 2-22　泰山刻石（局部）

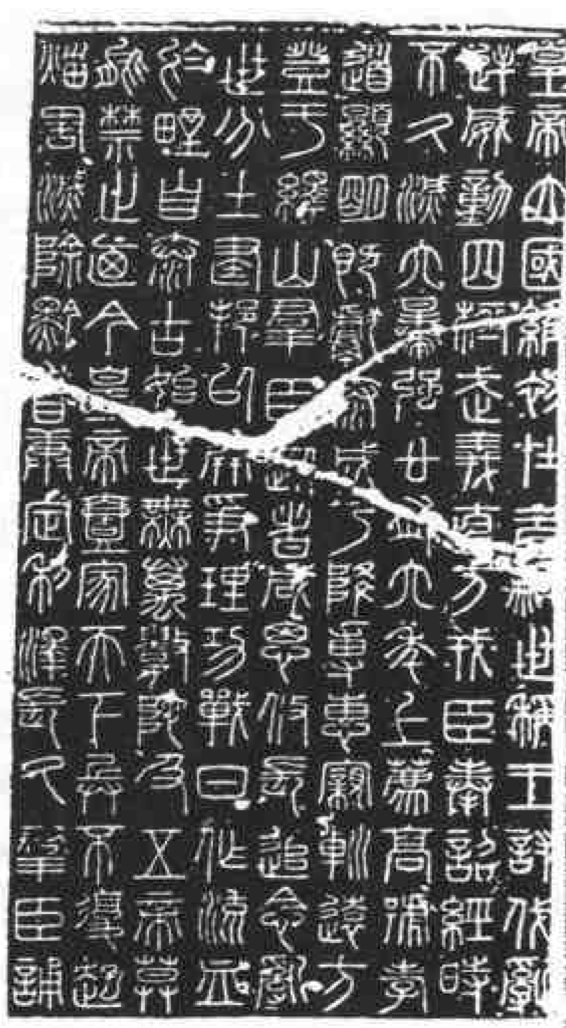


图 2-23　峄山刻石（局部）

立性很强，很难跟其他形声字建立联系，字际关系显得散乱而没有系统。到了小篆阶段，象形字在参与构字时大部分已经义化；形声字的表义部件的类化过程也已基本完成，代表同类事物的形声字多数采用了同一表义部件。这样，字与字之间的联系加强了，整个汉字体系显得系统而有条理。如：



“鸡”、“凤”在甲骨文中最初各像其形，添加声符变成形声字后仍是如此，二字之间在此阶段无法建立联系。但在小篆中，二字的象形部分均被“鸟”代替，这样，不仅此二字之间有了联系，而且还跟其他从“鸟”的字建立了联系，从而成为整个汉字构形系统的有机组成部分。《说文解字》关于540个部首的归纳，正体现了小篆构形的系统性。

（2）小篆比较全面地保存了汉字的构形理据。小篆是承西周的正统文字而来的，它虽然对古文字阶段的汉字形体作了系统的规整，却没有破坏汉字的构形理据。如“鸡”“凤”二字，经过类化以后，虽然字形已不能反映其实际形象，但其表义部件“鸟”仍能体现二字的物类特征，字的构形仍具有可解释性；而且这种系统化的理据比原先那种孤立的理据更符合汉字发展的要求。

（3）形声字大量增加。甲骨文中，形声字只占20%左右；而在《说文解字》中，则猛增至80%。不少原为象形、会意字的，通过添加形符或声符变成了形声字，如“鸡”、“凤”等；有些就干脆另造形声字来取代原来的象形或会意字，如：

沫：（甲）——→（小篆）　　箙：（甲）——→（小篆）

“沫”本像人洗脸之状，“箙”本像盛箭器，后来都变成了形声字。

（4）字形固定，异构、异写字较少。《说文解字》虽然收了一些“或体”字，但数量很少。大部分字的构件、构件的位置、构件的笔画都已确定，不能随意变动。这与甲骨文字形书写的高度自由形成了鲜明的对比。

（5）象形性减弱，符号性增强，笔画彻底线条化。《说文解字》说：“篆，引书也。”“引”就是牵引拉长，就是要使蜿曲的笔道平直化。“篆”书这个名称正反映了小篆逐渐脱离“随体诘诎”

的象形意味而走向线条化的特点。小篆不仅线条粗细均匀，笔画分布均衡，而且字的大小基本相同，字型均呈长圆态势，已初具方块汉字的规模。

总之，小篆是汉字发展史上第一次大规模的文字规范化运动，它一举结束了春秋战国以来汉字长期混乱的局面，为促进民族的团结和国家的统一作出了重大贡献。小篆是古文字阶段的最后一站，它对古文字形体的整理和规范，为汉字顺利过渡到今文字奠定了坚实的基础。

第三节　今文字阶段的汉字

一、隶　书

隶书是起源于战国晚期、到汉代趋于成熟的一种新型字体。隶书有秦隶、汉隶之分。秦隶又叫古隶，汉隶又叫今隶。

关于隶书的起源，《说文解字·叙》这样解释说：在秦始皇推行小篆的时候，“秦烧灭经书，涤除旧典，大发隶卒，兴役戍，官狱职务繁，初有隶书，以趣约易”。这种说法对错参半，隶书确实是因为职务繁多、需要提高书写速度而产生的，但并不起源于秦朝统一之后。早在战国时期的秦系文字里，隶书就已经出现。当时秦国的正统文字是小篆，笔道圆转回环，某些笔顺不符合手写的自然习惯，严重影响了书写的速度。而社会事务的日益繁杂，对快速书写的需求越来越迫切。于是书写人为了简便快捷，往往将篆文的圆转笔道改为方折笔画，有时还略加省改，已略具隶书的风味，在当时形成了一种俗书体。秦朝统一后，虽然将小篆定为标准文字，但日常事务中，隶书却得到了广泛的应用。1975 年在

湖北云梦县睡虎地出土的大量秦简，就是很好的证明。

秦隶实际上是篆文的较草率的速写体，它只是在笔形和态势上改变篆文的面貌，在结构上并没有太大的区别。从图 2-24 可以看出，秦隶明显处于篆文向汉隶的过渡阶段。

隶书发展到汉初，能够迅速取代小篆成为正统的字体，主要有三方面的原因：第一，汉字发展到小篆阶段，已形成了相当完善的构形体系；形声字的迅速增加，使得小篆总字数已近一万，完全可以满足当时记录汉语的需要；而且形声以其半形半声的构形特点，同汉语建立了音义双重联系，大大缓解了汉字与汉语之间的矛盾。这样，创造新字、调整内部结构已不再是汉字的首要任务，而小篆难写与社会对快速书写的需求之间的矛盾则上升至主要地位。在这种情况下，人们选择了较易书写的隶书，而抛弃了难写的篆书，这是符合汉字发展的自身规律的，因为汉字符号既要求具有区别律，也要求具有简易律，只有这样，才能更好充当记录语言的工具。第二，秦朝是个短命的王朝，他刚刚把小篆推上正统地位，自己就覆灭了，这就使得小篆失去了存在的政治基础。而新上台的汉朝统治者，多数是靠农民起义起家的，除了屠狗贩布的文盲，就是亭长、掾史之类的基层小吏，他们对隶书的熟悉程度远比小篆高得多。因而他们上台后用隶而不用篆，是情理中事。第三，小篆对古文字的整理和规范，也为隶书的迅速普及奠定了基础。经过小篆的整理

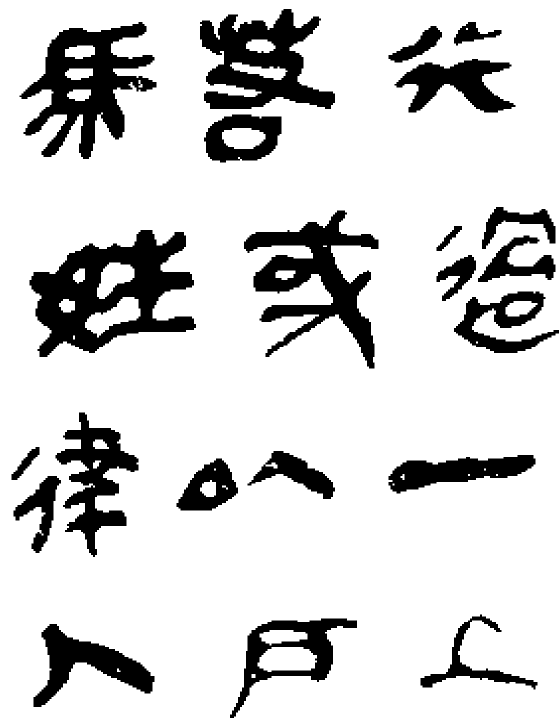


图 2-24 （转引自王凤阳《汉字学》）

和规范，汉字有了统一的形体标准，而当时的秦隶还明显带有篆书的余味，小篆的规范无形中也为秦隶提供了参照。隶书和小篆的这种密切关系，使得原先使用小篆的人在改用隶书时并不感到陌生。由于以上几种原因，进入汉代后，虽然没有明令废除小篆，但在日常应用中却很少有人使用，以致于东汉时期的一些读书人竟忘了曾通行过小篆这回事，“称秦之隶书为仓颉时书”（许慎《说文解字·叙》），甚至根据隶书字形乱解汉字，说什么“马头人为长”、“人持十为斗”等等。当时的经学家许慎对此极为不满，便写了《说文解字》以正邪说。

汉隶直接承秦隶而来，最初与秦隶并没有多大区别。经过二百多年的发展，到东汉时期，汉隶逐渐形成了自己独特的风格。笔势发扬舒展，笔画蚕头雁尾，有波势挑法；字体扁方平整，布局稳重匀称。如东汉中期的《史晨碑》（图 2-25）就是此期汉隶的代表。汉隶这种独特风格的形成，标志着隶书已经走向成熟。



图 2-25　东汉《史晨碑》（局部）

成熟的隶书在结构和态势上已与小篆有了很大的不同，这主要表现在以下几个方面：

（1）在书写笔法上将小篆圆匀的线条改成点画，从而彻底实现了汉字的笔画化。在此之前，汉字的书写单位是一些圆转的曲线，甚至是事物的整体轮廓，很难从中区分出笔画来，因

而我们前面一直称之为笔道。到了汉隶，汉字的书写单位基本上可以概括为一、丨、丿、㇏、丶等几种类型，真正意义的“笔画”才算形成，汉字笔画的计算也才开始成为可能。汉隶的彻底笔画化，使小篆所保留的那点象形意味丧失殆尽，汉字的符号性从此占据了绝对地位。这种变化，也使得汉字的书写难度大大降低，书写速度大大提高。

（2）汉隶对小篆的构成作了全面的调整。这次调整不象小篆那样是为了使汉字体系更为系统、更为严密，而主要是在简化动力的驱使下进行的。所以，在调整的过程中，更多的是照顾书写的方便和字体的布局，而不是照顾构形的理据。因简化而造成的偏旁合并、分化及省变现象，使得小篆原有的表意特征变得十分模糊。如

合并例：

①部首合并：

（肉部）　（月部）　（舟部）　（冃部）

肠　朗　服　胄

②成组部件的合并：

春　春　泰　奉　奏

分化例：

①没有规律的分化：

兵　莽　丞　奂

②因部首位置不同的分化：

（在左）　（在下）　（在上）

持　掌　看

（在左）　（在下）　（在右）

沙　泰　休

省变例：

→並　→雷　→冬　→泉

→曹→曹→曹

有时为了适应结构的平衡和笔画的匀称，隶书不惜把一个偏旁割裂开来。如“徒”字小篆作，从辵土声。因右部笔画太少，显得不太匀称，隶书便把徒分成彳和止两部分，将彳放在左边，占三分之一的位置；止置于右部土下，占三分之二的位置。这样，就把“徒”字纳入了“彳”部字的分布格局，从而取得了结构上的平衡。

(3) 在字体的态势上变小篆的长圆体而为扁方体，笔画故意向左右两侧取势，与纤细秀气的小篆相比，显得更为沉稳有力。试比较上举各例，就可明显看出这一点。

从篆书到隶书的转变，是汉字史上的一大飞跃，从此，汉字完全失去了古文字阶段的象形意味，摆脱了古文字蜿曲线条的束缚，开始步入今文字阶段。正因为这次转变给汉字带来如此剧烈的变化，文字学上便专门为它取了个名称，叫“隶变”。“隶变”对汉字的简化，是符合汉字发展的自身规律的。但它对汉字形义关系的破坏，给我们了解汉字的构意带来了很大的困难。

二、楷　书

楷书本名真书或正书，后因其结构严谨，堪为楷模，故改称楷书。

楷书萌芽于东汉，流行于魏晋南北朝，完全成熟于隋唐，一

直沿用至今。楷书产生的动因，同样是为了书写的方便。因为隶书的蚕头雁尾和波势挑法，仍然是影响书写速度的障碍。楷书则彻底摆脱了隶书的笔法，形成了标准的笔画，书写起来更加方便。

楷书直接从汉隶发展而来，它同汉隶在结构上基本相同，只是稍有简省。二者的区别主要表现在笔法和字体的态势上。楷书变隶书的扁方字体为正方，显得刚正典雅，端庄大方（图 2-26）。楷书的笔画与汉隶有明显的不同：横笔改为收锋，不再上挑；撇改为尖斜向下；钩是硬钩，不用慢弯；另外还增加了斜钩“㇂”（汉隶形如捺）、挑“㇀”（汉隶是横画斜写）、折“㇕”（汉隶是横画与竖画的自然结合）等基本笔画，这样，书写今文字所需要的各种点画已全部形成。所以说，虽然“隶变”已实现了汉字的彻底笔画化，但基本笔画的标准样式到楷书阶段才算最后定型。

从楷书成熟直到现在的一千多年中，汉字再也没有产生新的字体，平常所说的“欧体”、“颜体”、“柳体”、“赵体”，都只是楷书的不同书写风格，是书法意义上的“字体”，而不是文字学意义上的字体。解放后所推广的《汉字简化方案》，也只是对部分笔画繁多的汉字加以省改，并没有形成新的字体。



图 2-26　唐欧阳询《醴泉铭》

汉字字体停滞于楷书阶段，主要是因为汉字的构字笔画自定型之后，没再出现新的发展。另外，唐宋以后雕版印刷的盛行，使楷书形成了固定的模式，从而有力约束了汉字形体的变化。

第四节　汉字发展的一般规律

如果把新石器时代仰韶文化早期定为汉字起源的上限，那么，汉字已经有六千年的发展历史。在漫长的岁月和汉字所经历的实际变化中，我们可以看到汉字发展的几个规律和趋势：

（1）汉字在表意与表音的相互促进中，一直顽强地坚持自己的表意特点，不断地采用新的方式，增强个体符形和整个符号系统的表意功能。这主要表现在两个方面：

第一，当意义发生变化或符形笔势化以后，改造自己的符形和对字义的解释，以创造形义统一的新局面。例如，当汉字所记录的词所指的事物发生变化后，汉字终是及时调整它的义符，使之适应事物的特点。“砲”本是用石头做进攻武器，所以字从“石”。火药发明后，形体演变为从“火”的“炮”。“快”本从“马”作“駃”，交通发达后，马不再代表最高速度，“快”又引申出“快乐”义，形体便演变为从“心”的“快”等等。再如，当原初造字意图因符形演变而淡化时，汉字便采取添加意符的方式来增强其表音功能。象形字加义符的如“纹”、“韭”等等，会意或形声字累义符的如“捋”、“援”、“歌”等。这些都说明，汉字最终要最大限度地在符形上增加意义信息，来坚持自己的表意特点。

第二，由于书面语与口语互相转化（口语记录，书面语读出或唱诵），语音信息一时之间会变得异常强烈，加之有些抽象意义的词因义构形又比较困难，因此汉字在甲骨文时代就产生了一批

同音假借字，即，不再为某些词构形而借用同音字为符号。例如："戚"的本义是"斧子"而借作"悲戚"字，"舍"的本义是"房舍"，而借作"舍 "字，"介"的本意是"甲介"而借作"疆介"字等等。但是，表意的汉字对这些脱离意义的音化符号有"排异"作用，为时不久，它们便加上了相关的义符，分别写作"慽"、"捨"、"界"等，再度义化。正因为如此，汉字中的假借字向形声字转化，成为一种规律，形声字的声符是作为别词手段而存在的。

(2) 汉字在易写与易识的矛盾中，不断对个体符形进行调整，以实现简繁适度的优化造型。

汉字职能的发挥，是两个不可缺少的环节合成的，这就是书写和认识。就书写而言，人们终是希望符号简单易写；而就认识而言，人们又希望符号丰满易识。然而越简化，就越易丢掉信息，给识别带来困难；追求信息量大、区别度大，又难免增加符形的繁度，给记录增加负担。二者的要求是矛盾的。汉字就在二者的矛盾中相互调节，以追求简繁适度的造型。从甲骨文演变到小篆，随时都可以看到这种矛盾与调节。例如，("围"的古字)画四双脚围绕一个域邑，形象丰满，表意度高，但书写则不便，改为两脚作"韋"，意义又不很明确。最后加"□"作"圍"，不失本义而简繁适度。

在不影响表意与别词的前提下，汉字总是不断省去多余的分符与笔画，来减少书写的困难和减轻记忆的负荷。例如，大篆里，很多从"艸"的字是从"茻"的，小篆则一律减成"艸"，"集"曾写三个"隹"，到小篆里，将这些多余的重叠都减省了。可见，有限度的简化，是汉字发展的趋势。

(3) 汉字在发展中不断完善和简化自己的构形体系。

表意文字所遇到的最难解决的问题，是随着词汇不断丰富、意义不断增多，字形便会无限增加，致使符形量超过人有限的记忆能力。为了解决这个问题，汉字必须在对构件进行规整的前提下，

形成一个严密的构形系统。许慎的《说文解字》在人为调节的基础上，第一次把这个构形系统整理和描写出来。从小篆起，汉字就有了一批兼有意义和声音的成字构件作为构形的基础。许慎将小篆分为五百四十部，部首就是他规整出的基础构件。实际上，我们还可以进一步归纳为四百多个不重复的、音义具备的成字构件。其他汉字不论有多少，都是由这四百多个构件充当义符或声符按一定规律拼合而成的。汉字有了这批构件，便有条件把甲骨文的多形符象形字改造为多义符拼合的会意字，并且产生了一大批意符和声符相互制约的形声字。凡同义的字，用声符别词，如“根”、“枝”、“条”、“标”……都是树的部位，都从“木”，而用声符来区别，“玩”、“完”、“冠”、“顽”、“翫”……都从“元”声；而用义符来区别。这种形声字，大约占汉字总数的 87%以上，成为汉字的主体。同时，汉字的义符和声符的表意示源功能又都得到了进一步规整，这样，汉字便形成了以形声系统为中心的构形体系，由于采用了基础构件拼合生成的方法来增加新的字形，因此，不论字数如何增加，基础构件的数目都能保持稳定，只在四百多个上下浮动。而且，就构字功能说，构字量较大的基础构件只占 48%左右，人的记忆负荷是完全可以承受的。

形声系统形成后，汉字的表意方式发生了质的变化，人们对字符的认识不再是与物象直接联系的，也就是说，不需要从字形中直接辨识出物象来，而是凭借形音义已经结合了的基础构件来概括表意。例如：表示心脏的“心”发展为“[illegible]”，又发展为“忄”，早已不像心脏的样子，“[illegible]”发展为“氵”，也已没有水纹的痕迹。但“忄”形与“心脏”义，“氵”形与“水流”意，都已形成固定的联系，“心”部字与“水”部字的意义都可以由此辨识了。

汉字的构形系统形成后，仍然不断进行规整和简化。例如，基础构件既然不与物象联系，笔画便可以统一地进行合理的减省，隶

书的心演变为楷书的“忄”，就属这类减省。同功能的义符还可以合并，例如小篆的“鳥”与“隹”到楷书里作形符时渐渐合并为“鸟”，“佳”只作声符等等。声符的一部分在不破坏同源系统的情况下，还可按其示源作用进行规整，例如从“弘”与“宏”得声的字，已渐渐规整为从“厷”或简化为“厶”，从“宛”与从“盌”得声的字则规整为直接从“夗”，这样也可减省一部分声符。汉字构形系统的严密与简化是同时实现的，这是汉字发展的历史趋势中最重要的一点。

（4）汉字必须在自发发展的基础上进行人为的规范。

汉字的使用与发展都带有社会性，每一个汉字字符的创造与改变，一般都经过三个阶段：个人使用、社会通行和权威规范。从个人使用到社会通行，这是汉字自行发展的阶段；从社会通行到权威规范，这是人为规范的阶段。没有前一个阶段，汉字的社会性能便要丧失，它记录汉语的使用价值便会减弱，而没有后一个阶段，汉字演变的合理性和构形功能系统的形成规整和严密化是无论如何难以实现的。所以，汉字需要人为规范，又必须在尊重汉字的社会通行情况的基础上进行规范，在掌握汉字发展趋势并因势利导的前提下加以规范，这种规范才是具有社会性和科学性的。新中国成立以来所作的异体字规范和简化汉字的工作，从总的方向看，是符合汉字历史发展趋势的，因此，它才能被广大汉字的使用者所接受，收到最大限度的社会流通效果。

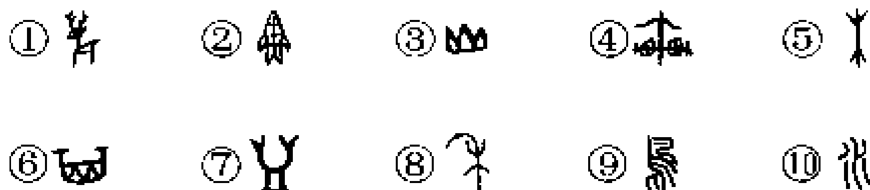
在汉字规范工作中，推行汉语拼音方案是十分重要的。汉字的形声字虽然占有很大比例，但形声字的声符本来就是用近似的声音来起别词作用的，并非准确的标音符号，何况经过数千年的历史演变，声符对形声字的直接标音作用更加弱化。作为表意文字的汉字，必须有一套科学的记音符号来协助标音，所以，推行汉语拼音方案，是弥补表意汉字不足的一个不可缺少的措施。

第三章
汉字的构形与构意

第一节　汉字因义构形的特点

汉字是表意文字，早期的汉字是因义而构形的，也就是说，汉字依据它所记录的汉语语素的意义来构形，所以词义和据词而造的字形是统一的。这一点，在小篆以前的古文字中表现得更为直接、明显。

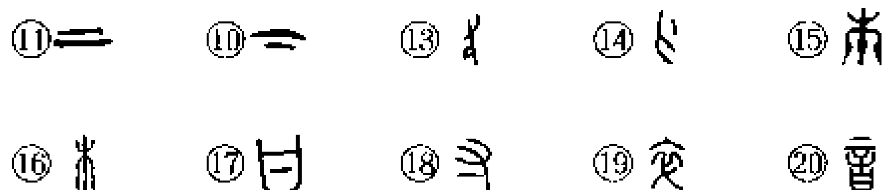
下面是一批古文字，有甲骨文、金文，也有《说文解字》中的小篆。可以看到，这些字的形体都是依据汉语的某一词义或语素义构造的，因此，这些字的形体也都可以用相应的词义或语素义来解释：



①～④分别是甲骨文的“鹿”、“鱼”、“山”和“车”。“鹿”像鹿形，以多歧的角为其特征；“鱼”直接描绘鱼的形象；“山”像山峰并立之形；“车”像车之两轮及辕、衡等形。⑤～⑧分别是金文的“木”、“其”、“皿”和“禾”。“木”上像枝，中像干，下像

根；“其”是簸箕的形象；“皿”像盛食器具之形；“禾”像禾苗之形，上像禾穗与叶，下像茎与根。⑨～⑩分别是小篆的“鸟”和“水”。“鸟”是鸟的象形；“水”是水纹的形象，因与“川”区别而将两边的线条断开。

①～⑩这10个字直接描绘物象，它们都与物象相似，一看就知道画的是什么，可以直接据形识义。



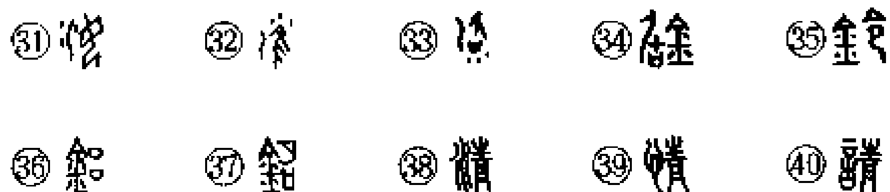
⑪～⑭分别是甲骨文的“上”、“下”、“厷”和“刃”。“上”和“下”是相对的概念，甲骨文的“上”和“下”分别用一根短线条放在一根长线条的上面或下而，以标识“上”或“下”的意思；“厷”像于臂肘上加标示符号以表臂肘之义；“刃”是在“刀”上加一短竖，指出刀刃的位置。⑮～⑯分别是金文的“本”和“末”。“本”是在“木”的下部加标示符号，表示下部的树根，“末”是在“木”的上部加标示符号，表示上部的树梢。⑰～⑳分别是小篆的“甘”、“寸”、“卒”和“音”。“甘”是在口中加标示符号，表示甘美的滋味；“寸”是在人手下寸口的地方加一标示符号，表示寸口；“卒”指士卒，古代士卒的衣服是有某种标记的，“卒”即在“衣”字上加指示符号，用士卒所穿的有标志的衣服来表示“士卒”的意义；“音”是在“言”上加“一”，以与“言”区别。

⑪～⑳这10个字要表现的是比较抽象的事物或难以直接画出的物体，只能用辅助的办法指出它们的位置或关系。虽然这些字需要一些略微抽象的思考，但它们也是属于因形而知义的。



㉑～㉔分别是甲骨文的“益”、“监”、“宿”和“采”。“益”像水在器皿中满溢的样子，这个意义后来写作“溢”；“监”像一个人用眼睛往盛水的器皿里看，古人以水为镜，“监”描绘的正是古人照镜子的形象，这个意思后来写作“鉴”；“宿”像人在屋中坐卧于席上之形，表示“止息”的意思；“采”像在木上采摘的形象。㉕～㉘分别是金文的“牧”、“取”、“初”和“集”。“牧”从“攴”从“牛”，以会“放牧”之义；“取”表示“获取”，《周礼》曰“获者取左耳”，故用以手取耳会意；“初”的意思是“裁衣之始”，故用“刀”和“衣”会意；“集”以众鸟集于木上表示“会集”义。㉙～㉚分别是小篆的“析”和“友”。“析”用以“斤”析“木”表示“剖析”义；“友”以两手相助表示“友善”之义。

㉑～㉚这10个字都是由两个以上的构件组成的。它们的意义由组成它的构件的意义、构件的数量及构件组合的相互位置来体现，字形和字义之间也是有统一关系的。



㉛～㉝分别为甲骨文的“汝”、“淮”、“洛”。它们都有相同的构件“水”，说明这些字的意义都与“水”有关，“女”、“隹”、“各”分别提示它们的读音。㉞～㊲分别为金文的“锯”、“铃”、“铝”和“铭”。它们共同的构件为“金”，说明这些字都与“金属”有关，它们的另一个构件“居”、“令”、“吕”和“名”分别

用来提示它们的读音。㊳～㊵是小篆的“清”、“情”和“请”。这三个字右边的声符相同，它们的区别是靠左边的义符来实现的，表示这三个读音与“青”相近的字，分别与“水”、“心理活动”和“言语”有关。

很显然，这一组字的字形和字义的关系不像以上三组字那样具体，它们的表义构件只是表示这些字字义的意义类别，也就是说，这些字的表义构件，已将它们的意义限定在某一个义类中，同时又有示音的构件表示读音。它们是在更概括的情况下实现形义统一的。

古汉字的形义统一是一个历史概念，其统一的情况存在着历史阶段的不同。

汉字的形体和意义都随着语言和社会的发展而处于不断发展变化之中，并且，在很多时候两者的发展变化不是同步的。有一些字，到一定阶段，形义关系就会变得不清晰，甚至脱节。但是，由于汉字是表意文字，据形识义是汉民族的文化心理，所以对于后代形义关系不清或形义关系脱离的字，人们又总是试图对发展后的字形和字义作出新的诠释，以期使形义关系得到新的统一。例如，《说文》小篆是比甲骨文晚、且又经过了规范的文字，许慎根据小篆的形体来分析构意，从而使形义关系统一，他的说解有很多是与甲骨文不同的。例如：

至，甲骨文作“ ”，画的是一支箭射到地面上，表达“到达”这个意义。小篆字形写作“ ”，《说文》对字形的解释是“鸟飞从高下至地也”，把上半部看成鸟展翅飞翔的形状。

东，甲骨文字形作“ ”，画的是一个两面扎口的口袋，就是后来的“橐”字。因为“东（橐）”的古音与“东方”的“东”相近，所以借用来表示“东方”的“东”。到小篆中，这个字的字形演变为“ ”，《说文》根据这个字形以及它与“杲”、“杳”等相关字形的联系，把它讲解为“日在木中”，表示太阳升起的方位，

即“东方”。

现代汉字都已变成了笔画，不再是具体物象的描摹，但由于现代汉字是从古汉字发展来的，与古汉字的形体是一脉相承的，因此大部分现代汉字的构意仍是可以讲解的。例如：

众：三人表示众多；

灭：火上有覆盖物则熄灭；

州：川表示水流，点表示水中的陆地；

意：存在心里的意义发出来就是音，所以意从音从心；

灾：屋内起火，表示成灾。

这些都说明，汉字作为表义文字体系，因义构形的特点是很清楚的，字形携带的意义信息也一直是十分丰富的。

第二节 汉字的书写单位和构形单位

一、汉字的书写单位

1. 笔画及其属性

汉字是刻成、铸成或写成的，刻纹、铸迹、线条和笔画是汉字的书写单位。

甲骨文是刻在龟甲、兽骨上的，它的书写单位是刻纹。大部分金文是铸在钟鼎上的，它的书写单位是铸迹。刻纹和铸迹都很难归纳出类型，也很难数出数目、定出刻铸的先后次序。到小篆阶段，书写的线条已相当规范，线条的样式已可分出类别，构成一个字的线条数也基本上可以数出来了。小篆的线条依形状可分为八类：

线条名称	定　　义	书写样式	举例
横	无曲、无折、无断的左右平放的线条	一	井　三
竖	无曲、无折、无断的上下直放的线条	丨	十　丄
斜	向左下或右下的直线	／　＼	[illegible]
弧	沿同一方向旋转而不封闭的线条	[illegible]	[illegible]
曲	中途方向改变的弧线	[illegible]	[illegible]
折	中途方向改变的直线	[illegible]	[illegible]
封	首尾相接的线条	[illegible]	[illegible]
点（断）	不足构成上述笔画的短线条	[illegible]	[illegible]

多数小篆的线条也可以数出数目，如：

大，3笔　[illegible]，3笔　[illegible]，4笔　[illegible]，8笔　[illegible]，8笔

但由于小篆图画性尚未完全淡化，因此线条书写的先后次序很难规范，笔顺没有形成。尽管如此，这已为隶、楷笔画的形成打下了良好的基础。

笔画是指汉字在书写过程中有始有终完成的最小的连笔单位。隶变以后，汉字线条的类型基本确定，一个字的线条数和书写顺序也可以确定了，也就是说，笔画形成了。到了楷书阶段，汉字的笔画就更为规范了。

笔画包括笔形、笔顺、笔画数等属性。笔形指笔画的具体形

状，笔顺指书写每个汉字时的次序和方向，笔画数指组成一个汉字的笔画的数目。

现代汉字的笔形一般可分为六大类：

笔画名称	定　义	书写样式	举例
横	无曲、无折、无断的左右平放的笔画	一	平王
竖	无曲、无折、无断的上下直放的笔画	丨	川竖
撇	向左下的斜笔	丿	八杉
捺	向右下的斜笔	㇏	大文
折	方向变化的连笔	㇕	凸及
点	构不成横、竖、撇、捺、折的短笔	丶	兴淡

这六类笔形是从汉字书写过程中的很多笔形中归纳出来的，实际上，每类笔形都包含很多变体。一般把“挑”看作“横”的变体，把“竖钩”看作“竖”的变体，因为大部分的“挑”和“竖钩”都是“横”和“竖”变来的，是“横”和“竖”为了和其他笔画的配合，由书写连笔造成的。也有人把“点”归入“捺”中，因为很多“点”是“捺”在与其他构件组合时的变体。笔形的分类可粗可细，如果为了检索、排序、教学或信息处理则宜粗；如果为了教授书法或描述写法则宜细，例如点可细分为横点、竖点、撇点、捺点、折点、提点、顿点等，小学课本就把汉字的笔画分为 24 类，但无论分得多细，也是以上述六类为基础的。

汉字在长期的书写过程中，为了书写的方便和间架结构摆布的美观需要，逐渐形成了先上后下、先左后右、先外后内最后封口的书写笔顺。与小篆比较，很简单的字形，小篆的顺序也难以固定，例如：

“甘”，小篆可以先弧后横：凵，甘；也可以先横后弧：—，

𠙴。而楷书则已形成了“丨，冂，口”的有规律的笔顺。

笔形和笔顺定型后，组成一个汉字的笔画数就可以确定了。汉字的笔画数是一种重要的检字手段，也是衡量汉字繁简程度的一个标准。

2. 笔画的组合

笔画与笔画的组合有以下三种情况：

（1）相离。例如：二、门、八、小、川

（2）相接。例如：丁、上、人、厂、斤

（3）相交。例如：十、七、力、并、也

笔画与部件的组合有以下三种情况：

（1）相离。例如：旦、引、术、孔、幻

（2）相接。例如：兀、天、灭、丕、夭

（3）相交。例如：夫、本、末、中、斥

把汉字的书写单位和构形单位区别开来，在理论上是非常必要的。首先，汉字的构件是体现构意的，笔画却不具有体现构意的功能。如“革”是以整体的构形来表示“去毛之皮”这一构意的，拆分成笔画后，各笔画体现不出构字意图。这就使构件与笔画有了根本性质上的差别。其次，汉字的书写与结构并不完全一致。书写是一笔一笔实现的，但不都是写完一个构件、再写第二个构件。只是在书写完成后，才能看见全部构件的布局。例如：

“回”由“囗（wei）”和“口”两个构件组成，但书写时并不是先写完“囗”再写“口”，也不是先写完“口”再写“囗”。

“東”由“木”和“日”两个构件组成，但书写时并不是将“木”和“日”一先一后分别实现的。

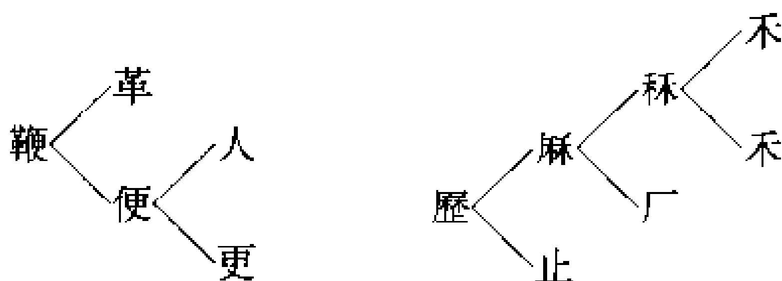
所以，虽然不少形素是由多个笔画构成的，我们在作构形分析时，并不以笔画作为下一层次的单位。当然，也有少数构件是单笔画的，为了理论体系的严谨，这种单笔构件应具双重身份：在书写时，称为笔画；进入构形时，称为单笔构件。

二、汉字的构形单位

汉字的构形单位是构件。当一个形体被用来构造其它的字，成为所构字的一部分时，我们称之为所构字的构件。如“人、止”是“企”的构件，“木、叔”是“椒”的构件，“冖、元、寸”是“冠”的构件。

（一）汉字的基础构形元素——形素

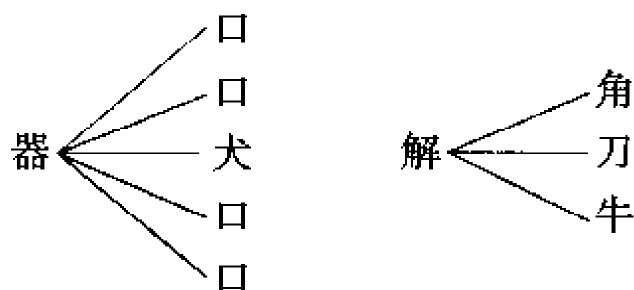
构件是有层次的，最小的基础构件我们称作形素。当我们对每个汉字进行拆分，拆到不能再拆，而又能够体现构字意图的最小单元时，这些最小的单元就是形素。例如我们对“鞭”字和繁体字“歷”进行拆分，依层次可拆成如下形体：



“革”、“人”、“更”就是楷书“鞭”和繁体字“歷”的形素；“禾”、“禾”、“厂”、“止”就是“歷”的形素。

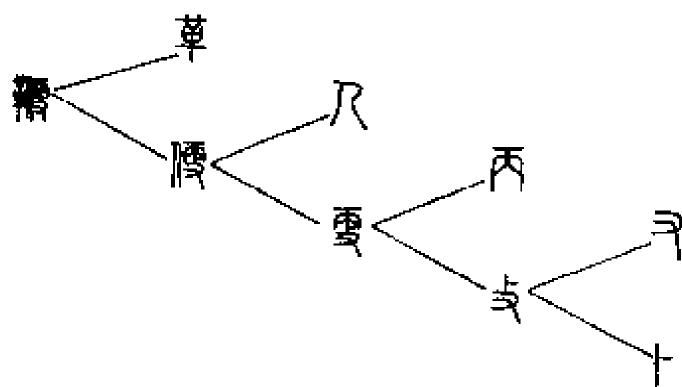
“鞭”和“歷”都是依层次拆分的，而下面的“器”和“解”则是一次性拆分的：

“鞭”、“歷”和“器”、“解”尽管拆分方式不同，但它们同样都是拆到不能再拆的程度，才现出了形素。这些形素都能体现一定的构意，例如“歷”中的“止”与“经历”义相关，“解”中的“刀”与“分解”义有关。大部分构件是由多个笔画构成的，我们在作构形分析时，却不拆分到笔画。因为笔画只是汉字的书写单位，而不是汉字的构形单位。如：“口”是以整体的构形来表示人



的嘴，拆分成笔画后，体现不出完整的构意了。

同一个汉字在不同的历史阶段，常会有不同的构形形素。例如，小篆“鞭”的形素与现代汉字就有所区别：



与小篆比较可以看出，现代汉字的拆分和构意的体现不如古文字细，这是因为从小篆发展到隶书、楷书，有时为了书写的便捷会发生形素的粘连，如“更”是“丙”和“攴”的粘连，“攴”是“又”和“卜”的粘连。粘连以后的构形元素从形体上不宜再拆分了。

（二）构件的层级

在依层次拆分的汉字中，构件是有层级的。例如：

“歷”字含有以下三级构件：

一级构件：止、厤

二级构件：厂、秝

三级构件：禾、禾

小篆的“鞭”字含有以下四级构件：

一级构件：革、便

二级构件：人、更

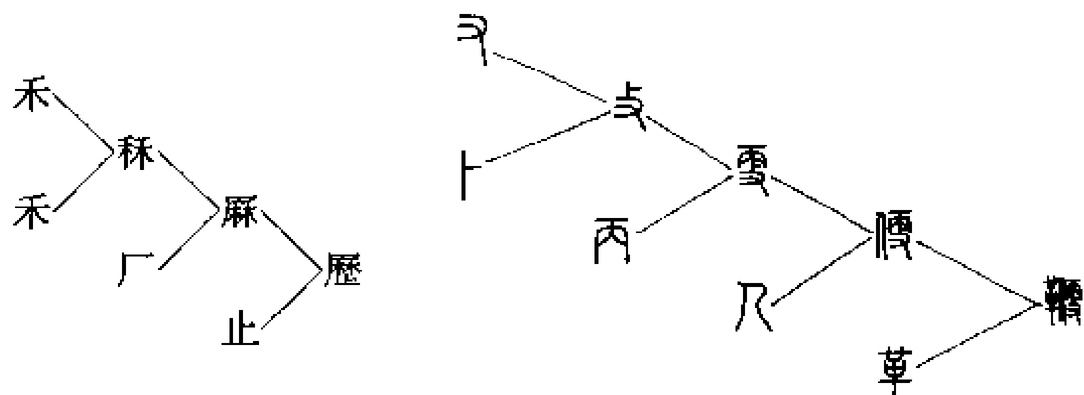
三级构件：攴、丙

四级构件：又、卜

一级构件又称直接构件，这样称说是因为它们直接构成全字，全字的造字意图是通过直接构件来体现的。如“鞭”的造字意图是通过直接构件“革”和“便”来体现的，“革”指明其意义类别，“便”提示其声音信息。最后一级构件即是形素，又称基础构件。

对独体字来说，它的构件就是它自己，它的基础构件、直接构件和全字是相同的，如“鸟”字的形素是“鸟”，直接构件和全字也是“鸟”。

以上我们是按拆分的程序来说的。将拆分的程序反转过来就是组合的程序，也就是汉字依层次逐级构成的顺序。楷书“歷”和小篆“鞭”的层次构成为：



从上可以看出，形素在层次组合的字中，是逐层加入的。也就是说，虽同是构形的最小单位，它们却可以有不同的出现层次。“又”和“卜”出现在组合的第一层，而“革”却到第四层才出现。

（三）成字构件和非字构件

构件有成字构件与非字构件之分。成字构件即能独立体现构意的构件。也就是说，当它不作其他字的构件时，本身就是一个完整的字，与语言中的某个词对应。如“日”，在作“晴”字的构件时，表示所构字的意义与“光亮”有关，而“日”本身就是一个独立的字，与语言中“太阳”这个意义相对应。“青”在作“晴”的构件时，其构意是提示“晴”字的读音，而“青”本身就有“qing”的读音和“青色”这个意义。“日”和“青”都是成字构件。

非字构件是依附于其它构件来体现构意的构件。这种构件本身不能独立存在，无法与语言中的词对应。如“果”中的“田”依附于“木”而存在，体现“果实之形”这一构意，它本身不能独立存在，不能与语言中的词对应。“太”字中的一点是为了和“大”相区别的一个标示符号，只能依附于成字构件“大”而存在，其构意只有在所构字的具体环境中才能体现出来。

有一种对汉字结构的传统分析方法，就是根据构件是否成字，将汉字分成独体字、准独体字和合体字三个大类：独体字是由一个成字形素构成的，如“人”、“口”；准独体字是由一个成字形素和一个非字形素构成的，如“果”、“刃”；合体字是由两个以上的成字形素构成的，如“息”、“伟”。

（四）构件在组构中的功能

构件在构字时都体现一定的构意，构件所承担的构意类别，称为这个构件的功能。汉字构件的功能有以下四类：

1. 表形功能

构件用与物象相似的形体来体现构意，我们就说构件具有表形功能。例如：

甲骨文的“（眉）”字中表示眉毛的构件“”，“（面）”字中表示面部轮廓的构件“”；金文的“（胄）”字中表示头盔的构件“”，“（旦）”字中表示日影之形的构件“”；小篆的“（果）”字中表示果实的构件“”，“（番）”字中表示爪印的构件“”等，承担的都是表形功能。

2. 表义功能

构件以它在独用时所记录的词义来体现构意，就是构件的表义功能。如：

“木”本指树木，在由它参构的字“杜”、“梅”、“桃”、“桐”等字中，“木”表示这些字的意义与“树木”有关。“言”本指言语，在由它参构的“説”、“話”、“談”、“語”等字中，“言”的构意是“言语”。“木”和“言”所承担的就是表义功能。

有些表形构件是非字构件，本身没有独立的音义，它们与表义构件的区分是明显的。但有些表形构件是成字构件，本身具有独立的音义，这样表形功能与表义功能有时就难以区分。我们可以从两方面来进行区分：

（1）表形构件是以象物性来体现构意的，因此只存在于古文字中。今文字的表形构件笔画化，原来的表形构件失去了象物性，一部分意化，成为表义构件，如古文字的“目”和“牛”就是这样；一部分记号化，必须用古文字的形体替代后，才能看出原来的造字意图，如小篆的“”和“”中表示果实和兽蹄的圆形物到楷书中变成了记号。因此，表形构件在今文字中已不存在。

（2）应当观察这个字的构件组合是形合还是意合。形合是指字的构件以象物性以及组合时上下左右的相对位置来反映事物的实际关系，从而体现构意。表形构件一般存在于形合字中。如“（涉）”字将两脚放在“水”的两边，以两脚跨越水的具体形象体现“涉水”之义，这样组合完全体现了事物的本来情境，所

以“涉”中的三个构件承担的都是表形功能。再如金文的“ (盥)”字，两手在器皿两旁，器皿中放了水，完全是洗手的实际物象的反映。整个字的构形属形合，它的构件即是表形构件。意合是指构件通过固定地贮存在形体中的意义来体现构意，而不以形体的象物性及构件的摆布位置与实际物象的对应来体现构意。楷书“涉”字，将古文字中跨于水的两脚移至右侧写成“步”，将“水”置于左侧，变成以“水”和“步”的意义信息体现构意的会意字，其构件都成了表义构件。楷书“盥”字中的所有构件虽还保留着原来的相对位置，但构件都已失去了表形作用，只提供“盥”与这些意义信息有关，其构件已成为表义构件。

3. 示音功能

构件在构字时体现示音的构意，就具有示音功能。例如：

“江”、“河”、“渭”、“洛”、“汾”、“湘”、“浙”是不同水的名字，都从“水”，构意与“水”有关，它们的构件“工”、“可”、“胃”、“各”、“分”、“相”、“折”分别与全字的读音相同或相近，承担的是示音功能。通过这些示音构件，可以将“江”、“河”、“渭”、“洛”、“汾”、“湘”、“浙”所对应的水名提示出来，并与其它水名区别开来。

在具有示音功能的构件中，有一部分还同时可以提示词源意义，也就是具有示源功能。例如：

“昏”的意义是“黄昏”，当它作“婚”的示音构件时，同时也提示“婚”由“黄昏”而得名。再如：“巠”的意义是“水脉”，具有“细而长”的特点，由“巠”作声符而构成的一些形声字，如“頸”、“莖”、“徑”、“脛”、“經”等，其词义都具有“细而长”的特点。因为提示意义来源只是示音功能附带的功能，不能单独存在，所以不单分一类。

4. 标示功能

构件不独立存在，而是附加在另一个构件上，起区别和指示

作用，即具有标示功能。承担标示功能的构件一般为非字构件。例如：

甲骨文“ （左）”中的“ ”是标示臂肘之所在，“刃”中的“丶”用以标示刀刃之所在，“本”下边的一横标示树根的位置，它们起的都是标示作用。“玉”中的“丶”是用以和“王”相区别，“刁”中的“㇀”是为了和“刀”区别开来，这些标示构件的功能是用以区别。

第三节　汉字的结构

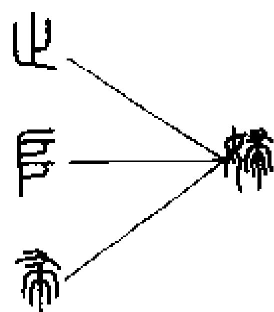
一、平面结构和层次结构

汉字由有限的形素组成数以万计的单字，有两种不同的组合类型，我们分别称之为平面结构和层次结构。

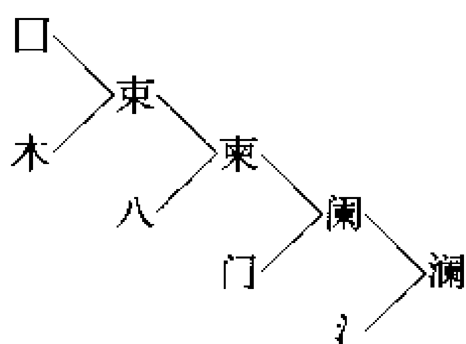
平面结构是由三个以上的构件一次性集合而成。前面谈到的“器”字和“解”字即为平面结构。甲骨文的“ （叟）”是“搜”的本字，像手持火炬在室内搜索之形，它的三个构件“ ”、“ ”、“ ”是一次性集合而成，分不出组合的先后次序。甲骨文的“ （薅）”像手持农具除草之形，它的四个构件“ ”、“ ”、“ （辰：一种农具，即蚌镰）”和“ ”也是同时参构的。小篆“ （归）”是“女嫁”的意思，它的三个构件“ （止）”、“ （妇）”和“ （自）”也是一次性集合而成，分不出组合的先后层次。

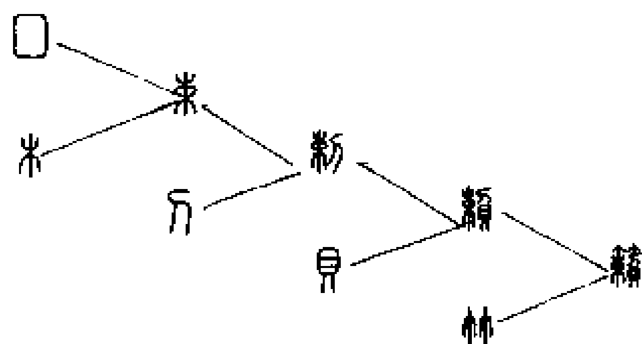
这种平面结构的字，其构意的体现有两种类型：一种是把多个意义信息集在一起来表示总的字义：如小篆的“归”由“止”和

“帚（妇）”提示意义，“自”提示读音；楷书“解”用“牛”、“角”、“刀”三个表义构件一起体现构意。这些字都已不是实际物象的组合，它们是通过构件本身所带有的意义信息或声音信息来体现构意的。另一种是图画式的组合：甲骨文的“叟”和“薅”就是通过对实际物象的描绘来传达“搜寻”和“除草”之义。



但是，汉字的大量组合却是层次组合，全字是由形素或构件分作若干层次逐步累加上去而构成的。以小篆的“籁”和楷书的“澜”为例：





这两个字都是由形素逐次累加而成的。这种层次结构的汉字，其构意是通过直接构件来体现的，其他构件不直接对全字的构意起作用，只是因为逐级生成全字而对全字的构意间接地起作用：

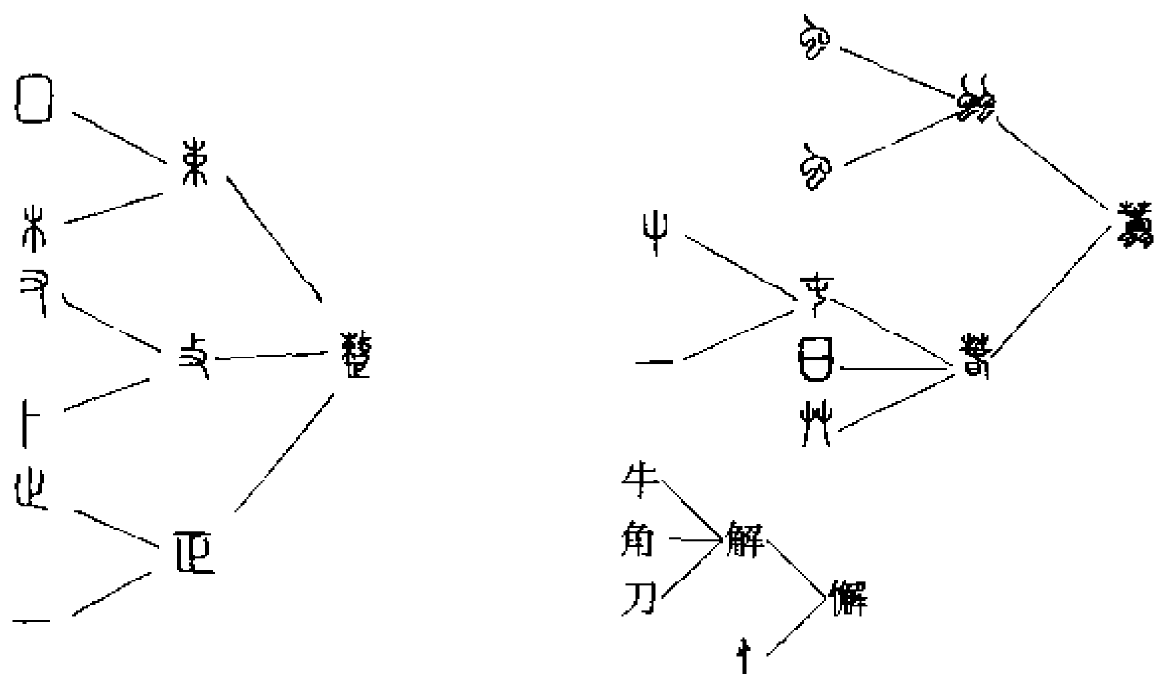
籁：由“口”和“木”作为组合的起点，生成“束”，表示“束缚”义。“束”和“刀”组合为“刺”，产生了“乖戾”义，徐锴说：“束而乖违者，莫若刀也。”在下一个层次里，“刺”以声符的身份和义符“贝”组成“赖”，既保持了“刺”的声音，又产生了“赢利”的意义。“赖”又作为声符和“竹”组合构成“籁”，表示乐器的名称。

澜：先由“口”和“木”生成表示“束缚”义的“束”。“束”和表示“分别”义的“八”构成“柬”，表示“挑拣”。

“柬”作为声符和“门”组成“阑”，生成“门前栅栏”的意思。“阑”又作为声符和“水”组合为“澜”，既保留了“阑”的声音，又产生了“大波”的意思。

由以上的例子可以看出，在层次组合中，形素的功能只在自己加入的层次里起作用，每一个层次都生成一些功能，也可能消失一些功能。整个字的构意则是通过直接构件来体现的，而直接构件的功能又是通过下层构件逐级生成的。

有一些字，属于平面组合和层次组合综合构成。如小篆的“整”、“蠢”和楷书的“懈”：



“蠢”是经过三个层次组合而成的字，在组合的第二层出现了平面组合。“整”和“懈”都是两层次构成的字，“整”的第二层是平面组合，“懈”的第一层是平面组合。

因为平面结构与层次结构在构意的体现上，前者是集合式的，后者是生成式的，所以在分析汉字的形体结构时，正确区分这两种结构类型，才能保证部件的拆分不出错误，也才能准确分析造字的理据。

二、汉字的构形模式

构件的形体在作平面组合和层次组合的过程中，是有功能参与的。构形模式是指构件以不同的功能组合为全字，从而体现构意的诸多样式。根据构件的组合情况和构件在组构中的功能，我们从历代汉字的实际状况出发，总结出 11 种构形模式：

1. 全功能零合成

由一个成字基础构件构成的字即为全功能零合成。因为只有一个成字构件构成，它既表义又示音，因此它对全字所起的作用是全功能的。这个构件从构件成为全字，在身份和功能方面都发生了变化，我们假设这种变化是由它和“零”合成而引起的，因此称这种构形模式为“零合成”。如：

甲骨文的“ （萬）”字，金文的“ （马）”字，小篆的“ （龟）”字等，全字的形、音、义信息都是由构件自身提供的。隶书和楷书中的“日”、“人”、“更”等已经失去了原来的表象物性，但音、义信息仍由它自身提供，也是零合成字。

2. 标形合成

由一个表形构件加上标示构件，即为标形合成。如：

小篆的“ （寸）”是以表形构件“又”加标示符号“一”构成，表示寸口所在的位置；“ （母）”是在表形构件“女”中加两点，标示哺乳女性的特征；“ （甘）”是在“口”中加标示符号，表示甘美的滋味，它们的构形模式都是标形合成。

3. 标义合成

由一个表义构件加标示构件，即组成标义合成。例如：

小篆的“ （音）”是由表义构件“言”和标示构件“一”构成；楷书的“吏”由表义构件“史”和标示构件“一”构成；

"太"由表义构件"大"和标示构件"、"构成。

4. 会形合成

两个以上的表形构件组合在一起，即为会形合成。会形合成字都是以形合的方式来组合，也就是说，这种合成字不但构件是以物象体现意义，而且按物象的实际状态来放置构件。如：

甲骨文的"（龋）"由表形构件"（齿）"和"（虫）"构成，以齿中有虫的组合方式表示龋齿；"（贮）"由表形构件"（贝）"和"（宁）"构成，且将"贝"置于构件""之中，形象地传达"贮藏"义。

一般说来，这种会形合成字如果构件在三个以上，大多是平面组合。例如：

甲骨文的"（降）"，以两个向下的"止"与""（象山形）同时组合，表示向下走；"（陟）"以两个向上的"止"与""组合，表示向上走。小篆的"（莫）"由三个表形构件""、""、""及其相对位置关系表示"日暮"，都是以事物的实际状态为依据来平面组构成字。

5. 形义合成

用表义与表形构件组合在一起，表示一个新的意义，即为形义合成。例如：

小篆的"（胃）"由表形构件""和表义构件"肉"构成。再如，"柬"的本义是"挑拣"，小篆作""，由表义构件"八"和表形构件"束"构成，"八"有"分别"义，"束"像枝叶被捆缚状，全字的构意是把捆在一起的东西分拣开来。

6. 会义合成

用两个以上的表义构件组合在一起，表示一个新的意义，即为会义合成。会义合成字的构意，是由表意构件所提供的诸多意义信息融合而成的。"尘"由表义构件"小"和"土"构成，"雀"由表义构件

“小”和“隹”构成，“逐”由表义构件“辶”和“豕”构成，“宝”由表义构件“宀”和“玉”构成，这些都是会义合成字。

以上的六类构形模式，都是表形构件、表义构件和标示构件组合，没有示音构件的介入。汉字有了示音构件的介入，字面所含的信息量更大了，而且有了更经济的区别手段，在汉字构形史上，是一大进步。

7. 标音合成

用标示构件与示音构件组合，即为标音合成。如：甲骨文的“𠂉（千）”是由示音构件“人”和标示构件“一”构成。“百”是在示音构件“白”上加标示构件“一”构成。标音合成字是用来区别同音字或近音字。因为汉字具有因义构形的特点，因此用标音合成构成的字不多。

8. 形音合成

用表形构件与示音构件组合，即为形音合成。这种模式在甲骨文中是一种很重要的模式。甲骨文中的一些象形字，出于区别或更便于识别的原因，再加上一个示音构件，以增加字音的信息，便成为形音合成字。例如甲骨文中的“凤”、“鸡”、“星”等，原来都是象形字，以后又增加了“凡”、“奚”、“生”这样的示音构件，使字面所含的信息更为丰满。再如，小篆的“齒”由表形构件“𠚕”和示音构件“止”构成。



9. 义音合成

用表义构件与示音构件组合，即为义音合成。它以表义构件来体现义类，又以示音构件来提示读音，形成了同类字以音区别，近音字以义区别的格局，例如：

①组：河　何　诃　蚵　坷　苛　岢　钶　珂　轲　疴　呵

②组：混　演　涣　滂　汪　浩　滥　洪　滔　涌　溃　淙

①组字音都与“可”相近，以“水”、“人”、“言”、“虫”、“土”、“艹”、“山”、“金”、“玉”、“车”、“疒”、“口”表示其义类而区别；②组字义都与“水流的状态”有关，以“水”充当它们共同的表意构件，而又以“昆”、“寅”、“奂”、“旁”、“王”、“告”、“监”、“共”、“舀”、“甬”、“贵”、“宗”提示其声音而区别。

下面再讲两种综合的结构模式。它们都是适合平面结构的构形模式：

10. 无音综合合成

这种合成是表形、表义与标示构件的一次合成，但没有示音构件介入。例如小篆的“[illegible]（葬）”，由表示草的表形构件“屮”、“屮”，表示“死”的表义构件“[illegible]”和标示构件“一”组合而成。

11. 有音综合合成

这种模式是由多个表形、示音、表义、标示构件一次合成的。例如甲骨文的“[illegible]（渔）”从两手举网，鱼声。甲骨文的“[illegible]（春）”从三个“木”，从“日”，“屯”声。

三、传统的汉字构形分析方法——“六书”

上面我们讲到的汉字构形分析方法，是在界定了部件的构意功能之后，根据直接部件组合的结构来确定每个汉字的模式的，也就是一种“结构－功能”分析法。这种分析方法与传统的“六书”是什么关系呢？

汉字发展到小篆时代，许慎将字形进行了一次大的规整，同时也将构形模式作了一次简化。在《说文解字》中，许慎提出了

分析汉字构形的“六书”，并给它们下了定义，举了例字：

> 一曰指事，指事者，视而可识，察而可见，上下是也；
> 二曰象形，象形者，画成其物，随体诘诎，日月是也；
> 三曰形声，形声者，以事为名，取譬相成，江河是也；
> 四曰会意，会意者，比类合宜，以见指㧑，武信是也；
> 五曰转注，转注者，建类一首，同意相受，考老是也；
> 六曰假借，假借者，本无其字，依声托事，令长是也。

严格说来，只有前四书（象形、指事、会意、形声）是用来说解形义关系的条例，转注和假借是用来说明文字在使用过程中字与字之间的关系的。

象形字即根据文字所要记录的词义，用直观构图或图形象征的手段构成文字的一种造字方法。如小篆的“門（门）”“衣（衣）”“冊（册）”等。

指事字要表现的是比较抽象的事物或难以直接画出的物体，只能用辅助的办法指出它们的位置或关系。它们或者直接用抽象符号指示出一种位置关系，如“丄”、“丅”等；或者是在原有的象形符号的基础上加上标示符号，如“刃”、“亦”和“寸”等。

会意字是利用已经造出的象形字或指事字进行拼和，来体现新的意义。例如“取”、“森”、“析”、“歪”、“闯”、“闪”等。

形声字是靠着一个表示义类的义符和一个提示声音的声符构成的。能够与意义联系的是义符，声符只是提示字的读音，帮助我们找到该字所记录的词。例如“烧”、“燎”、“烟”、“精”、“瞳”、“简”、“模”等。

这样看来，“六书”的前四书也是用“结构—功能”分析的原则确立的。

前面讲过的 11 种构形模式，大约可以涵盖自甲骨文以来的可

以分析构意的各类字形。我们知道，自今文字以来，构件就没有表形功能了，因此，有表形构件参与的模式也就随之消失，也就是说，11 种模式只剩下 7 种，这就是构形系统的简化。而系统简化的工作，从许慎就开始了。小篆经过许慎的规整，首先把构件成字化，也就是加以意化，表形构件一般只出现在独体字和准独体字中，把合体字构件的功能简化为表义、示音和标示三种，构形模式也就简化为前四书了。

前四书对经过许慎规整过的小篆字形是契合的，但是，它无法涵盖小篆以外的其他字体，特别是古文字字体。一般人勉强用前四书分析各类汉字时，常按小篆把独体字称作象形字，采用标示构件的字称作指事字，有示音构件的合体字都笼统称形声字，没有示音构件的合体字都笼统称会意字。

上述十一种构形模式和“六书”中的前四书，以及汉字结构的三分法，可以大致形成以下的对应格局：

<table>
<tr><td>全功能构件＋0</td><td>零合成字</td><td>独体字</td><td>象形</td></tr>
<tr><td>表形构件＋标示构件</td><td>标形合成字</td><td rowspan="3">准独体字</td><td rowspan="2">指事</td></tr>
<tr><td>表义构件＋标示构件</td><td>标义合成字</td></tr>
<tr><td>示音构件＋标示构件</td><td>标音合成字</td><td rowspan="4">形声</td></tr>
<tr><td>表形构件＋示音构件</td><td>形音合成字</td><td rowspan="7">合体字</td></tr>
<tr><td>表义构件＋示音构件</td><td>义音合成字</td></tr>
<tr><td>示音构件＋各类构件</td><td>有音综合合成字</td></tr>
<tr><td>表形构件＋表形构件</td><td>会形合成字</td><td rowspan="4">会意</td></tr>
<tr><td>表形构件＋表义构件</td><td>形义合成字</td></tr>
<tr><td>表义构件＋表义构件</td><td>会义合成字</td></tr>
<tr><td>各类构件（无示音）</td><td>无音综合合成字</td></tr>
</table>

在用结构功能分析法把汉字分成11种模式以后，可以看出勉强用前四书分析各类字体的做法确实有削足适履的弊病。比如，前面已经说过，会形合成字与会义合成字在特点上是很不一样的，会形合成字的构件所居的位置是与构意有关的，如果在考据时把构件位置不同的会形合成字释为同一个字，势必要产生误识。又如，传统独体字在小篆里的确绝大部分还保留象物性，但是这些字在隶变、楷化以后已经义化，再称为象形字很不妥当。楷书中相当一部分独体字是粘合而成，如果用独体象形字的观念去看，更是难以理解了。再如，甲骨文里的标音合成字，用标示构件来分化同音字，它既不能进入形声字——因为其中没有表义或表形构件，又不都用成字构件构成；算作指事字也不妥当——因为它有示音的要素，可以说，传统"六书"里找不到它的位置。所以，我们现在用"结构－功能"分析法确立的构形模式，对各时代的汉字都可以囊括，它体现了"六书"的基本精神，避免了"六书"的局限，也能涵盖前四书，在理论上和操作上应当是可行的。不过，"六书"的影响实在太大了，在普及的领域里，有时还不得不使用它来大致界定汉字的构形、解释汉字的构意，但有了对结构功能分析的全面认识，对"六书"条例的使用，可以更慎重一些，力求不产生失误。

第四节　汉字构形与构意的分析

一、汉字形义统一的条件

在第一节讲过，因义构形是汉字的特点，但是，汉字又是处在不断的发展变化中的。形体、意义和使用三方面的发展变化，都

会造成形义关系的不清晰甚至脱节。当我们通过形体来证实意义，或通过意义来解释形体时，必须知道，汉字的形义统一是有一定条件的，这些条件主要有三个：就字形来说，必须用表示笔意的形体；就字义来说，必须用本义；就字用来说，必须用本字。

1. 笔　意

我们把还保持着原来的造字意图，可以从字面上直接观察构意的形体，叫做笔意化字形，简称“笔意”。第一节提到的那些携带着意义信息的甲骨文、金文、小篆和现代汉字的字形，都是可以据形识义的，我们说这些形体都保留了笔意，或者说，都是笔意化的字形。

汉字发展到今文字阶段，由于笔画的形成，形体趋于简化和整齐，出现了一些脱离了原始造字意图的形体，形义的统一关系不能直接显现出来，也就是发生了书写的笔势化，例如：

“亦”，小篆作“ ”，是在一个正面的人形的腋下加上标示符号，指示“腋窝”的意思。楷书把表示正面人形的“大”字笔画化，上下断开，写作“亣”，便失去了笔意。

“兼”，小篆作“ ”，像一只手握住两把禾苗，因面有“兼得”义。楷书改曲线为平直笔画，两个“禾”字的线条省并、变形，表示手的“又”字失去象形特征，构意就变得模糊了。

这些都说明，一部分楷书的形体与原初的构意已经不能吻合了。在讲解这些字形的意义时，字形与字理产生了矛盾。解决矛盾的办法是找到笔意。

2. 本　义

字形发展的同时，字义——也就是构形所依据的词义，也在不断发展变化，致使形义关系不清晰。我们所说的形义统一的“义”，必须是本义。本义是与字形相贴切的词的义项，是由笔意分析出来的，可以用文献来证实的词义。例如“木”的本义是“树木”，“监”的本义是“照镜子”，“益”的本义是“满溢”，

"宿"的本义是"止息"。

词义在应用过程中，会从本义出发，沿着本义的特点所决定的方向，按照民族心理和习惯，不断引申出很多新的意义，这就是引申义。如人们一般认为树木的知觉比较迟钝，因此"木"由"树木"义引申有"麻木"义。"照镜子"须仔细审视，故"监"从本义出发引申出"察看"义，再引申出"监牢"义。"益"由"满溢"引申有"多"、"增多"、"好"等意思。"宿"是夜间的行为，一宿过后即为第二天，因此"宿"就有"往昔"、"旧时"的意思，这就是"宿怨"、"宿仇"等词中的"旧时"义。

近引申义与字形的关系尚能间接看出，远引申义，如"监"的"监牢"义，与字形的关系则已变得模糊不清了。探讨字形与意义的关系，一定要用本义来讲，才能达到形义相统一，如果用引申义特别是远引申义讲字形，容易产生牵强附会的弊病。

3. 本　字

本字指的是专为记录某个词所造的字，它的形体是以这个词的某一义项为根据而造的。如：

"其"是"簸箕"的本字，故甲骨文的"其"像簸箕形；"汝"是"汝水"的本字，故字形从"水"；"亦"是"腋窝"的本字，故"亦"的古文字字形是用符号标示"腋窝"的位置。再如："信用"这个意义的本字是"信"，"伸展"这个意义的本字是"伸"。本字是据词义而造的，是形义统一的。

但文字在应用的过程中，同音借用的现象很多。如：

用"其"字记录第三人称；用"汝"字记录第二人称；把"亦"用为副词。这些地方用的就不是本字，而是因为读音相同而借用的字，我们称这种字为同音借用字，简称借音字；由于同音借用而使本字具有的意义称作借义。如："汝"在记录第二人称"你"这个意义时是一个借音字，第二人称"你"这个意义是"汝"的借义。再如，该写"伸"的地方写了"信"，如《孟子·告

子上》"今有无名之指，屈而不信"这句话中，"信"和"屈"相对，记录的是"伸展"的意思，这里的"信"是"伸"的借字，"伸展"义是"信"的借义。

借音字与其所记录的字义是脱节的，用借音字的形体来讲解它所记录的意义，必然会出错。只有本字才能保证形义关系的统一。

以上从字形、字义、字用三个方而说明了形义统一的条件，一言以蔽之，我们说的形义统一就是本字的笔意和本义的统一。

二、分析构意的基本方法

现代汉字虽然多数具有可分析的笔意，但在隶变、楷化的过程中，也确实出现了一批因粘合、省简、变形、错讹而变得看不出造字理据的字。在文献中，虽然多数时候用本字，但遇到借字的机会也非常多。而对字义来说，用引申义的机会比本义多。对于那些不具备形义统一条件的字，在通过构形分析构意，进而探讨文献词义的时候，或者通过词义分析汉字构形的时候，需要做三方面的工作：

1. 笔意的恢复

现代汉字构形理据丧失在字形方而的原因是笔意的模糊，这一点又是由以下几方而的原因造成的：

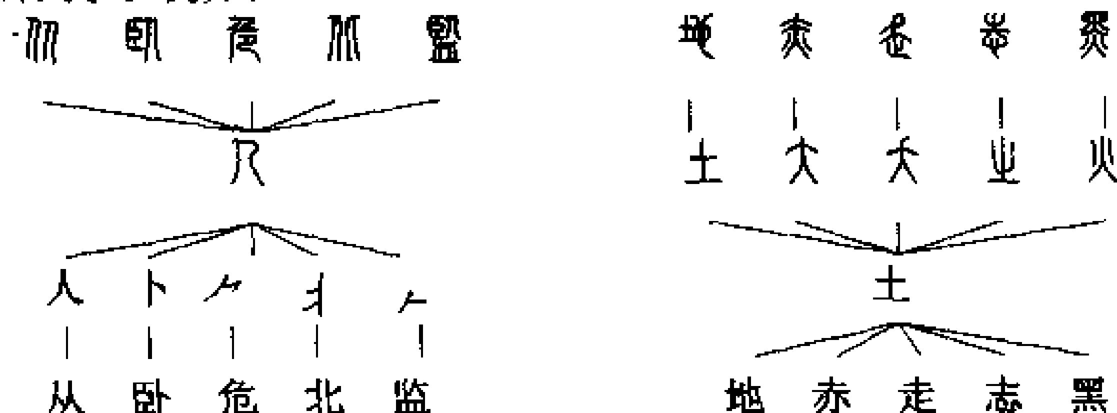
（1）隶变造成的象形性消失。古文字象物性很强，经过隶变，书写形制发生了由曲到直的变化，使字形的象物性淡化以至消失，构形理据不能直接由字形看出。例如：

"虫"小篆写作"[小篆字形]"，是个象形字，隶书写作"虫"，象物性消失。有人根据隶书字形把"虫"的构意胡乱解释为"虫者屈中。"

“矢”小篆作“ 𠆧 ”，象箭形，隶书作“矢”，失去了象物性。

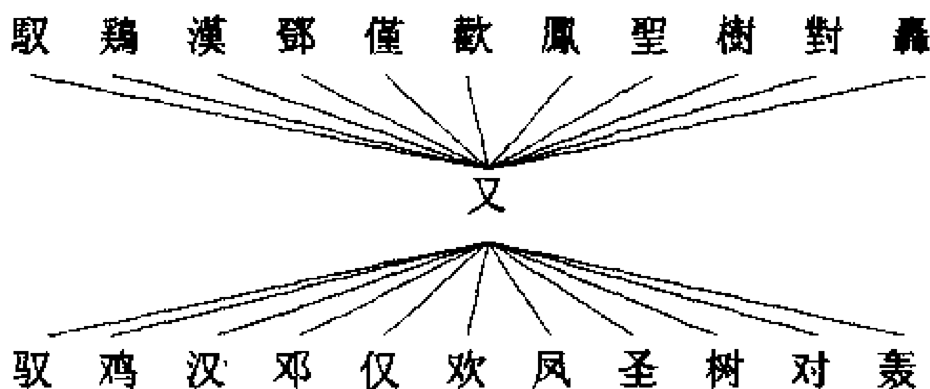
“衣”小篆作“ 衣 ”，是上衣的象形，隶书作“衣”，失去了象物性。

(2) 隶变和现代汉字简化造成的形体变异与混同。原来的一个形体变成了几种不同的写法，又使得很多原来不同的形体变得没有了区别，这样，好多字的构形理据就不能通过字形的分析而得到了。例如：



隶变使“人”变成了“卜”、“⺈”、“丬”等，但“人”的构形理据是绝不可以用“卜”等字形来解释的。隶变使“大”、“夭”、“火”等不同的形体都变得和“土”没有了区别，但这些字的构形理据也是不可以用“土”来解释的。

在简化现代汉字时，符号替代造成一些部件合并。如：



以上的字中，除了“驭”外，其它字中的“又”字已经完全变成了人为约定的记号，这些字的构形理据与这些记号没有任何

联系。

（3）隶变造成构件的粘连，使构形理据无法从字面获得。如：

“更”在小篆中写作“ ”，从“攴”、“丙”声，经过隶变“丙”和“攴”粘连不可分了，其构形理据也就不能直接从字形看出了。

“寒”在小篆中写作“ ”，人藏在室内草中，草下有冰，表示寒冷，隶书作“ ”，中间的两个“艹”和“人”粘合在一起，构形理据无法从字面获得。

隶变和现代汉字简化造成字形象物性消失和构件的变异、混同、粘连，使构形理据变得模糊不清。要想根据字形讲解构意，就必须找到笔意化字形。通过变异字形寻找笔意的工作我们叫做“复形”。例如：

“为”有“劳作”义，但楷书“为”的字形无法与“劳作”联系起来，这就要对“为”字作“复形”的工作，即找到保留着原初造字意图的形体。我们找到“为”的甲骨文字形“ ”，是手牵大象进行劳作的形象，这个字形与“劳作”义是统一的。再如：

“昔”有“往日”的意思，楷书“昔”从“艹”从“日”，“日”可以和“往日”联系起来，但上部的“艹”是什么？通过复形，我们找到“昔”的甲骨文字形“ ”，上像洪水。传说古人曾经历过一个洪水时期，那次洪水给人类带来了深重的灾难，也给人类留下了难忘的印象，所以古人用“水”和“日”会意，表示“往昔”。通过复形找到字形来源，构形理据就可解释了。

汉字由起源而渐进为现代楷书，其间经历了很多的发展阶段。就形体的外观来说，不同历史时期的汉字存在较大差别，但各历史阶段又是相互联系、一脉相承的，汉字字体的演变又是有一定内在规律和一定发展趋势的。我们在找笔意的时候，就要依照汉字形体演变的规律和趋势，只有这样，才能做到不主观，才能保

证科学。

2. 本义的探求

很多现代汉字的常用义都已不是造字时所依据的本义，而是后出的引申义，致使形义不能统一。我们在作构形理据的分析时，要注意找它们的本义。因为本义乃是引申义的源头，是统帅一系列引申义的义项，在寻求引申义和字形的关系时，必须通过本义作过渡。如：

“企图”义是“企”的引申义，这个引申义与字形的关系已经不是那么直接了。“企”的本义是“踮着脚看”，这个意义与“企”字从“人”从“止”的构形是相合的。

“横”在“横眉冷对千夫指，俯首甘为孺子牛”中有“逆”、“不顺”义。这是“横”的引申义，我们要想弄明白“横”字何以有“不顺”义，就得通过“横”的本义作考察。《说文·六上·木部》：“横，阑木也。”阑木就是挡门的横木，它拦住开门的方向，阻挡门的推进，所以引申有“不顺”、“横逆”的意思。由“不顺”又引申为“错综杂乱”的意思，由此我们才能更清楚地理解诸如“横生枝节”、“沧海横流”、“妙趣横生”、“才华横溢”、“血肉横飞”等这些词语中“横”的意义。而“横”有“与理相悖”的意思，用在“横行霸道”、“横征暴敛”等词语中，也就不难理解了。

古代文献中，“尉”有“安慰”义，如：

《汉书·韩安国传》：“以尉士大夫心。”颜师古注：“尉，古尉安之字正如此。其后流俗乃加心耳。”根据颜师古的注我们知道，“尉”本身就有“安慰”义，后来才加了表义构件“心”作“慰”，来专表“安慰”义。“尉”何以有“安慰”义？《说文·十上·火部》：“尉，从上案下也。从𡰥，又持火以尉申缯也。”“尉”即今天“熨斗”的“熨”的本字，“尉”的目的和结果都是使一个东西平展，引申而有“安慰”义。

3. 本字的追溯

汉字在使用过程中，借音字的存在使形义关系脱离，要想使形义关系统一，必须追溯其本字。追溯本字必须首先明确两点：第一，只有形体无法与意义统一时，才需要求本字，而求到本字后，形义便应当是统一的；第二，求本字的目的不是为了写本字，而是为了从本字的形体上探求或证实它所记录的词义，从而使形义统一，加深我们对文献词义的理解。本字虽是专为某词而造的字，但在使用时，有时反而是借字通行而本字不通行。我们在使用文字符号时，对已经通行的借字要遵循社会约定俗成的原则。如：

“举国欢腾”、“举世闻名”的“举”的本字应为“俱”，《说文·八上·人部》：“俱，偕也。”引申而有“全部”义。但“举”已经通行，我们不必改为本字。

“雕刻”的“雕”的本字是“彫”，《说文·九上·彡部》：“彫，琢文也，从彡周声。”但“雕”已通行，大家都这样写，并没有造成交流的障碍，也不必用本字来取代借字。

如何通过借字找本字呢？首先通过声音来找，因为借字就是和本字声音相同或相近才被借用的。例如：

“容”，《说文解字》：“容，盛也。从宀谷。”“宀”和“谷”都是容积很大的空间，“容”的本义是容纳。“容貌”的“容”和“容纳”义无关，我们怀疑是借字。根据声音关系，我们认为其本字应是“颂”。“颂”的异体字作“額”，从“容”得声，证明“颂”与“容”在古代是同音的。“颂”从“页”，字义与头部有关，《说文》：“颂，貌也。”“颂”正是“容貌”之“容”的本字。

“衣冠楚楚”是说一个人穿得很好，“楚楚动人”是说一个人的长相非常美好。可是《说文解字·木部》对“楚”的解释是：“楚，丛木，一名荆也。”“楚”是一种植物，也就是今天所说的“荆棘”的“荆”这种植物，和衣物无涉。根据声音关系，我们找到一个“黼”字，《说文·七下·黹部》：“黼，合五彩鲜貌。从黹

虘声。《诗》曰"'衣裳髗髗'。"《说文》是用来说明本字本义的，从《说文》对字义的解释和所引《诗经》我们知道，"衣冠楚楚"之"楚"的本字正是"髗"。

同音借用问题涉及到许多疑难词义的探求和解释，是一个十分重要的问题。但通过借字找本字仅仅有语音上的相同或相近还是不够的。因为，汉字的同音字很多，从古到今时代久远，语音变化的轨迹也往往曲折绵长，如果妄加猜测而找错本字，便会得出十分荒谬的结论。由借字找本字需要从文献上找到充足的证据，不容主观臆断。例如：

成语"不速之客"的"速"，注释家一般用"召"来训释它，认为不速之客为不招而自来的客人。如《周易·需卦》有"有不速之客三人来"之语，王弼注："不待召也。"孔颖达解释说："不须召唤之客有三人自来。"二者皆以"召（即今天的'招'）"训"速"。"速"字从"辵"，本义是"疾速"，不可能发展出"召"的意思来，我们怀疑这是一个借字。根据音同音近的条件，我们找到一个"誎"字。《说文》："誎，餔旋促也。""餔"是下午四五点钟吃饭，有客不按时来，恐天黑不便，常需促请。"旋促"就是催其快来入席。"誎"由本义"催客人吃饭"引申而有"请"、"召"之义，所以"不速之客"的"速"，本字应是"誎"。为了证明这个本字的推断，还需从文献中多找一些有关同音借用字的佐证。例如，《诗·小雅·伐木》中有："即有肥羜，以速诸父"；"既有肥牡，以速诸舅"的句子，都是借"速"为"誎"的用例。

在汉语词汇中，有些同音借用字已经通行，而且并不一定有本字，对这样的字，不必硬找本字，也不要硬讲构意。例如：

"离"的繁体字作"離"，《说文·四上·隹部》："離黄，鸣则蚕生。""離"从"隹"，它的本义是一种鸟的名字，意义与"分离"义无涉，用来表示"分离"，是一个借字，但后来"離"成了"分离"义的通行字，本字已经不明。

第四章
汉字的构形关系
和使用关系

第一节　汉字的构形关系

汉字的构形关系包括两种情况：一种是共时的，即在同一个历史时期的汉字之间的形体关系；另一种是历时的，即在不同历史时期汉字构形的传承关系。共时的构形关系包括异写字、异构字，历时的构形关系包括传承字。另外还有一种特殊的文字现象，就是同形字。

一、异　写　字

异写字是同一个字（音、义、用完全相同的字）因写法不同而造成的形体差异。

在早期汉字中，由于书写带有随意性，形体不固定，所以异写现象大量存在。例如：甲骨文的“羽”有多种写法：



再如金文的“𠂢”，像水脉之形，也有很多种写法：



异写字主要有两种：一种是构件写法变异造成的异写字；一种是构件位置不固定造成的异写字。

1. 构件写法变异造成的异写字

如甲骨文的“虎”字是一个零合成字，它的许多写法就是由书写变异造成的：



再如：金文中的“酉”字有许多异写形体，当它充当其他字的构件时，又会造成所构字的许多异写形体，金文“奠”的许多

异写形体就是主要由"酉"的异写造成的：



到了汉字的使用非常广泛以后，在个人使用汉字这个层面上，由于书写人的个人习惯不同，也会形成一些异写字，例如隋唐碑志中就有许多异写字，下边是一组"龙"字的异写字：



不同历史时期的汉字积淀到同一共时层面上，隶定方式不同也造成异写字。如："春"从小篆发展到楷书，出现了两种不同的字形"春"和"萅"，"萅"是将小篆的字形用楷书的形制来转写的结果。在隋唐碑志中，"垂"有"垂"和"埀"两种写法，"光"有"光"和"灮"两种写法，也都是隶定方式不同造成的。

2. 构件位置不固定造成的异写字

在有些字中，构件放置的相对位置不同，可以形成不同的构意，如"東"、"杲"、"杳"等，在这些字中，构件的位置就成为字与字之间的区别手段。但是，也有相当一部分字，构件放置的位置是不影响构意的。这些构件位置不同而构意相同的字，也是异写字。在早期汉字里，这种构件位置不固定的异写字相当多。如甲骨文的"吏"字或作"　"，或作"　"，手的位置在左在右不定；"祀"字或作"示巳"，或作"巳示"，"示"和"巳"的位置

可以互换。金文的“福”，或作“福”，或作“[illegible]”，“示”和“畐”的位置不定等等。在汉字规范时取消的异体字中，有不少其实是这类异写字。例如：

案——桉　　鞍——鞌　　峰——峯　　鹅——鵞

濶——闊　　幕——幙　　棋——棊　　慼——慽

异写字的相互差别只是书写方面的差别，没有构意方面的差别，从社会用字的角度看，这是一种对全社会文化交流不利的现象。首先，它增加了汉字使用者的负担，同是一个字，却要记很多形体；其次，它影响印刷与汉字信息处理，使排版与打印字库增加许多无谓的形体，字形增加，信息量却没有增加；第三，在基础教育与对外汉语教学中，字形的混乱使教育者与受教育者备受干扰，文化的传播也必然因此受到影响。所以对异写字必须进行规范。规范异写字的方法有两种：

第一，直接在异写字中确定一个字形作标准体，其他形体只要能与标准体认同就可以了。例如在上各例中，确定“峰”、“鹅”、“幕”、“棋”为标准体，其他的形体，在教科书和公开的社会用字中，一律取消；遇到历史文献用字，只要与标准体认同，用标准体置换，也就达到规范的目的了。

第二，对基础构件加以规范。由上述各例可以看出，异写字的形体差异，很多都是在汉字最小的基础构件内部产生的，如果我们把汉字的基础构件形素规范化，由基础构件构成的字也就避免了异写现象。例如“奠”的异写字可由“酉”来规范，“龙”的异写字可由“[illegible]”来规范。

二、异　构　字

异构字是与异写字相对的一个概念。异构字是指在任何情况下音与义都相同，而在构件、构件数量、构件功能等方面起码有

一项存在差别的一组字。异构字的构意肯定是有或多或少的差别，这就是我们称之为异构字的原因。但异构字记录汉语的功能是相同的，它们在书写记录言语作品时，不论在什么语境下，都可以互相置换。异构字有以下两种基本类型：

1. 因构形模式不同而形成的异构字

①[illegible]——[illegible]

金文的“皿”是象形字。“⿰金皿”是在“皿”上加表义构件“金”，原来的“皿”就成了示音构件，“⿰金皿”就成了义音合成字。

②韭——韮

“韭”是象形字，小篆作“[illegible]”，《说文解字》：“象形，在一之上。一，地也。”“韮”是义音合成字。

③丘——坵

“丘”是象形字，甲骨文作“[illegible]”，像山丘形。“坵”是义音合成字。

④岳——嶽

“岳”从“山”从“丘”，是会义合成字。它的异构字“嶽”是从“山”、“狱”声的义音合成字。

⑤泪——淚

“泪”从“水”从“目”，是会义合成字。它的异构字“淚”从“氵”、“戾”声的义音合成字。

⑥埜——野

“埜”是从“林”从“土”的会义合成字。它的异构字“野”是从“里”、“予”声的义音合成字。

2. 构形模式相同，构件不同的异构字

（1）有些异构字都属于会义合成，而表义构件不同。如：

灾——災　明——朙　尘——塵　⿱死土——塟——葬

（2）有些异构字都是义音合成字，而表义构件不同。如：

鲠——骾　雞——鷄　醎——鹹　唇——脣

（3）有些异构字都是义音合成字，而示音构件不同。如：

线——線　　粽——糉　　岭——险　　咽——嚥

“俯”和“俛”等都是示音构件不同的异构字。

（4）有些异构字都是义音合成字，而表义构件和示音构件都不同。如：

村——邨　　寨——砦　　碗——盌　　腿——骽

从上述诸例可以看出，异构字之间的差别，是构件及构件功能等构形属性上的不同，因此不能将它们认同为一个字，只能说它们是记词功能相同、构形不同的两个或数个字样。异构字的差别既然发生在构件、构形模式和构形布局等构形属性上，它们的构意也就必然不相同。在利用字形探求词义的时候，多一个形体，便可以多一个认识角度，异构字在帮助我们对字所记录的词的本义有更清楚的认识方面，能起到很大的作用。

但是，异构字的记词职能是完全一致的，在记录语言时，是一种不必要的重复。这些字的存在不但增加了识字量，还有可能使读者视异字为异词，造成理解上的障碍，所以异构字在社会用字中应当加以规范。不过，规范异构字只能优选一个字形作为通行字，另外的形体在记词功能上认同，却不能贸然取消。因为这些字在某一历史层面上虽成为异构字，到另一历史层面上却有可能分别承担着几种不同的功能。因此，整理汉字时，遇到异构字要慎重对待。

三、传　承　字

传承字是指不同历史时期具有传承关系的汉字，可以说它们是不同时期、不同构形系统中的同一个字。我们把确定字与字之间的历史传承关系称作汉字的历时认同。证明字与字之间有历史传承关系，首先需要认同它们的记词职能，因为传承字记录的是

不同时期的同一个词。

历时汉字的传承大约有以下三种情况：

(1) 汉字经过发展，因为总的形制变化，笔法有所不同，但构形全同，全然可以断定是同一个字。例如：

小篆的“ ”用圆转线条；而楷书的“果”形制发生了变化，笔划变平直了，但构形仍是上像果形下从“木”。

甲骨文、金文、小篆和隶书、楷书的“日”字，笔势上虽有差异，构形属性却是一脉相承的。由于构形属性没有发生变化，构意的解释当然也不会发生变化。

(2) 汉字经过发展，结构要素、结构模式和结构层次与分布都没有改变，只是形素的样式发生了变化，直观上不像同一个字，但只要通过构形分析，便可知道它们是同一个字。例如：

甲骨文的“果”作“ ”，上面的果形比小篆的形体更具象形意味，但仍是上像果形下从木，这种变化是发生在结构元素的内部，发生在组合的最底层。

金文的“受”作 ，是用上下两手受付舟（承盘）来会授受之意，发展到小篆作“ ”，上面一只手略有变形，中间的“舟”简化为“冖”，下面一只手不变。

笼统地直观地看来，它们的形体变化很大，但采用形素分析法将变异的形素认同后，很容易分析出它们的形体传承关系。这类传承我们称作形体变异传承，简称变异传承。变异传承最值得注意的是一部分后来变成结构部首的汉字。例如：

小篆的“ ”，楷书写在右边时变为“阝”；

小篆的“ ”，楷书写在左面时变成“氵”；

小篆的“ ”，楷书写在左边时变成“衤”。

这种变异我们称作位移变异，它的特点是，传承前的汉字是成字，传承后成为非字，只能充当构件，不能单用了。但是，含

有这种变异构件的字构意没有变化，传承关系是显而易见的。

（3）汉字在发展过程中，发生了构形元素或其他构形属性的改变。例如：

金文的“射”，是以手持弓箭形，到小篆时，弓箭变成了“身”字，《说文解字》将其构意解释为：“弓弩发于身中于远也。”

[illegible]——[illegible]

甲骨文的“尽”从皿从手持涤器，像食毕洗涤器皿的形状，战国文字稍有变异。小篆因手持涤器形与手持毛笔的“聿”字相近，因而将字形上部改为从“聿”从“火”的“𦘔”字，整个字成为从“皿”、“𦘔”声的义音合成字：

[illegible]——[illegible]——[illegible]

“更”在周代金文中上像钟形，下边从“攴”，以示用锤击钟。小篆变为从丙声，从“攴”，到隶书时“丙”与“攴”粘合成一体，无法分析了：

[illegible]——[illegible]——[illegible]

由于构件或其他构形属性发生了变化，构意的解释自然也不同了，但是在形体上还能找到后一时代字形由前代字形演变过来的轨迹。

历时认同是古文字考据的基本功，也是现代文字溯源的基本功，在汉字构形学中，它的应用意义是很大的，但也是较有难度的，我们在这里只是大概言之。

四、同　形　字

同形字是指记录不同词的相同字形，并且几个词的词义没有引申或假借关系。上述异写字、异构字和传承字的记词功能相同，在文字整理和文献阅读时，需要作认同的工作；同形字则是不同

的词用了相同的形体，在文字整理和文献阅读时，则需要将同形字区分开来。

同形字从其产生方式上大致可分为两种：

1. 造字的同形

分别为不同的词造字，而形体偶然相同。如：

“杷”，从“木”、“巴”声，是收麦器的名称，今天写作“耙”。人们还造了一个“杷”字，记录“把手”的意思，今天写作“把”。这样当“收麦器”讲的“杷”和当“把手”讲的“杷”在造字的时候就同形了。

“铝”，周代金文里有一个“铝”字，指铸铜器的原料，与现在的金属铝无关；汉代扬雄的《方言》中也有一个“铝”字，义为雕磨；而现在的“铝”字则指一种金属。这三个“铝”字是不同时代的人为不同的事物而造的同形字。

我们把金文的“铝”和现在的“铝”看作同形字，是就它们积淀到楷书层面的字形来说的，也就是说，这些字在用楷书书写时变成了同形字，而就分别为它们造字的当时来说，它们属于不同的字体，形体并不相同。如金文的“铝”作“[illegible]”，和现代汉字的“铝”是不同形的，但当我们用楷书转写当时传下来的文献时，或在字书中用楷书转写古文字时，两者变成了同形字。

还有一些字，本来并不同形，后来因为重新造字或改换构件而变得同形了。如：

“忧”，《说文解字》有一个“忧”字，“忧，心动也。从心，尤声。”“忧愁”的“忧”本作“憂”，《说文解字》：“憂，和之行也。从夊，㥑声。”表示“悠闲”义。在简化现代汉字时，为“憂”造了一个新形声字“忧”，与表示“心动”义的“忧”同形了。

“腊”是一个形声字，读xī，本义为干肉。“臘”，本义为农历十二月举行的一种祭祀活动，故称农历十二月为“臘月”。近代民间把“臘”的声符写成“昔”，就与表示“干肉”的“腊”同形了。

“价”，古音 jiè。《诗经·大雅·板》：“价人维藩。”郑玄注释说：“价，甲也。被甲之人，谓卿士掌军事者。”依此说则“价”为甲士。价格之“价”（jià）原作“價”（初文作“贾”，本指商人，引申指价格），后简化作“价”，与“价”（jiè）同形。

“叶”，本音 xié，为古“协”字。《后汉书·律历志》：“遂覲东后，叶时月正日。”树叶的“叶”（yè）原作“葉”，后简化为“叶”，与“叶”（xié）同形。

“怀”是形声字，从“心”、“不”声，本义是“怒”。“懷”也是形声字，从“心”、“褱”声，本义是“思念”。后来人们把“懷”写作“怀”，与表“怒”义的“怀”同形了。

2. 书写演变的同形

为几个词所造的字本来不同形，在形体演变或书写的过程中变得同形了。如：

“胄”，在《说文》中本为两个字，一个在“肉”部，作“胄”，“胄，胤也。从肉，由声。”胤即有血统的后裔；一个在“冃”部，作“冑”，“冑，兜鍪也。从冃，由声。”兜鍪就是现在所说的甲胄。隶变以后，二字同形了。

“栝”，在《说文》中本为两个字，都在“木”部，一个作“桰”，“桰，檃也。从木，𠯑声。”即矫揉曲木的工具；一个作“栝”，“栝，炊灶木。从木，舌声。”即烧火棍。隶变以后，“𠯑”旁变成了“舌”，“桰”和“栝”变成了同形字。

在文字的书写过程中，因书写省减或书写随意性等原因也造成一部分同形字。如在隋唐碑志中，“木”作偏旁时，经常写作“扌”，这样就造成一些同形字：“標領”的“標”有时写作“摽”，就与“擗摽”的“摽”同形了。“楊柳”的“楊”有时写作“揚”，就与“飞揚”的“揚”同形了。“楷模”的“模”有时写作“摸”，“兰桂”的“桂”有时写作“挂”，“楷模”的“楷”有时写作“揩”，“樞纽”的“樞”有时写作“摳”，都造成了同形字。

由于同形字实际上是用一个字去表示两个或两个以上的词，在阅读文献时，我们就必须搞清该字在具体语境中究竟表示的是哪个词，否则，就会将不同意义的词因书写形式相同而认同为一个词，从而影响对文献的理解。

第二节　汉字的使用关系

一、同音借用字

文字承担着与它的字形相一致的本义，也承担着由本义引申而尚未分化出去的引申义，这是它固有的职能。但文字在使用时，其记词职能却常常发生变化，其中一个很重要的原因是有些文献在使用汉字时存在着因声音相同而互借的情况，也就是说，汉字在使用过程中，有时不写本字而写借字。遇到这种情况，本字的职务就转移到借字身上了。

借字代本字的现象，又分两种情况：

1. 造字阶段的假借

造字阶段的假借即许慎在《说文解字·序》中所说的"本无其字，依声托事"的"假借"。当语言中产生了某个新词来不及造字，或本应为这个词造字，但某些词词义抽象，很难据义构形（如副词、连词和表人称的词等），便借已有的同音字来表示这个词，不再造新形。这些字的形体，无法用后一词的词义来解释。例如：

"它"是"蛇"的本字，表第三人称的"它"未造字，而借本义为"蛇"的"它"字记录。

"夫"是"丈夫"义的本字，代词"夫"的词义难以构形，于

是借“丈夫”之“夫”为之。

“難”是从“隹”、“英”声的形声字，本义为鸟名。“困难”义难以构形，就借用“難”字来记录。

2. 使用阶段的假借

使用阶段的假借就是“本有其字的假借”，又分两种情况：

(1) 借字通行，本字废除。本来为某个词造了字，但本字未通行，借同音字记录，同音字反而通行了，本字表义的职务转移到借字上。如：

“健康”之“康”的本字应是“犺”，《说文·十上·犬部》：“犺，健犬也。”引申泛指“健康”。《广雅·释诂》：“犺，健也。”但“健犺”之“犺”常借“康”字为之。“康”即“糠”的本字，本义应是“谷皮”，但后来在表达“健康”之义时，“康”行而“犺”废。

“旧”的繁体作“舊”，《说文·四上·萑部》训“鸱舊”，是一种鸟名。“新旧”之“旧”的本字应为“肍”。《说文·四下·肉部》：“肍，熟肉酱也。从肉九声，读若旧。”肉酱必经久而成，故引申而有“陈旧”义。

“雅座”是指整洁、安静的客席。但“雅”字从“隹”，本义是一种鸟，《说文·四上·隹部》：“雅，楚鸟也。”“雅座”显然不应是“鸟座”，这个“雅”是“房”的借字。《说文·九下·广部》：“房，庑也。”“庑”是客堂，房座是设在客堂的座位，也就是待客之席。但今天“雅”行而“房”废。

“毛”的本义是动物或人身上的毫毛。“羊毛”、“鸿毛”、“毛发”等都用的是本字本义。但“毛手毛脚”、“毛躁”的“毛”是借字，本字应是“冒”，《说文·七下·冃部》：“冒，冢而前也。”意思是“没有目标地往前走。”正是“毛躁”、“毛手毛脚”的本字。但“毛”通行而“冒”不通行。

本字的职务转移到借字身上后，由于借字的长期沿用，约定

俗成，借字反而变成承担这个意义的通行字并为用字者和阅读者所承认。这种情况一般不影响对意义的理解，只有当我们想通过字形对字义作进一步的了解时，才需要通过声音线索找到本字，了解其形义结合的原始状态，从而弄清词义的发展轨迹。

(2). 临时代用，借字、本字同时通行。本有其字且本字通行，但有时不写本字，临时用其它同音字代替。例如：

①借“裁”代“才”　《促织》：“复之以掌，虚若无物；手裁举，则又超乎而跃。”这里的“裁”当“刚刚”讲，与“裁”的本义“剪裁”无关，“刚刚”的本字应是“才”。《说文·六上·才部》：“才，草木之初也。”引申为“刚才”、“刚刚”义。

②借“归”代“馈”　《论语·阳货》：“归孔子豚。”这里的“归”是借字，本字应是“馈”，即“馈赠”之义。

③借“涂”代“途”　《孟子·梁惠王上》：“涂有饿莩而不知发。”“涂”在这里是“道路”的意思，当为“途”的借字。

④借“蚤”代“早”　《孟子·离娄下》：“蚤起，施(读 yí)从良人之所之。”“蚤”是“早”的借字，在这里是“早晨”的意思。再如：《史记·项羽本纪》：“不可不蚤自来谢项王。”这里的“蚤”也是“早”的借字。

⑤借“弊”代“蔽”　《韩非子·定法》：“君无术则弊于上，臣无道则乱于下，此不可无一，皆帝王之具也。”“弊”是借字，本字应为“蔽”，在这里是“受蒙蔽”的意思。

⑥借“冯”代“凭”　《左传·僖公二十八年》：“请与君之士戏，君冯轼而观之。”“冯”是“凭”的借字，《说文·十四上·几部》：“凭，依几也。”“凭轼”就是依傍着车轼。

⑦借“窥”代“跬”　《中山狼传》：“君能除之，固当窥左足以效微劳，又岂讳之而不言哉？”这个“窥”用的是借字，本字当是“跬”，当“举足”讲。

这种临时代用字使形义关系脱离，往往影响对文意的理解。有

些虽经注疏家将意义注出，仍很难知道这个意义从何而来。因此，就必须沿着声音找到本字，从而做到对词义的正确而透彻的理解。

二、职能分化字

职能分化是指原来用一个字记录的词或义，为了区别，分开用两个或两个以上的字来记录。用来分化职能的字，就是职能分化字。文字符号是要求相互区别的，但是由于意义引申或同音字借用，汉字的记词职能常常兼用，如“臭”，从“犬”从“自”，本义是“嗅闻”，后引申指“气味”，这样“臭”就具有两个意义，即“臭”字兼有了两方面的职能。到一定的时候，从求区别的原则出发，就会产生字的分化。如，为了表义明确，后来在“臭”字上加表义构件“口”造了分化字“嗅”，专门记录“嗅闻”义。

分化文字职能的方法，主要有三种：

1. 借用同音字分化职能

（1）借用同音字承担本义：当假借义或引申义成为一个字的常用义后，其本义就借用另一个同音字来表示。例如：

“何”的本义是“担荷”，它的甲骨文字形作“ ”，画的就是一个人有所担荷之形。由于“何”被假借用来表示疑问代词“何”，并且这个借义成为它的常用义，因此，它的本义“担荷”就借“荷花”的“荷”来表示。

“内”的本义是“纳入”，《说文·五下·入部》：“内，入也。”《荀子·富国》：“婚姻娉内，送逆无礼。”用的就是“内”的本义。后来，“内”引申而为“内外”的“内”，这个引申义成了它的常用义，它的本义就假借“纳”字表示。“纳”字从“糸”，本义是形容湿丝的状态的，《说文·十三上·糸部》：“纳，丝湿纳纳也。”

（2）借用同音字承担假借义或引申义：当本义仍是本字的常用义时，则借用其他同音字来记录假借义或引申义。例如：

“女”本义指“女人”，曾假借作第二人称代词，后来，这一借义又改借“汝水”的“汝”来表示。

“指”本义是“指示”，引申而有“意旨”的意思，后来假借表示味美的“旨”来表示这一引申义。

借来分化职能的字，通常是本义跟假借义很少发生混淆的字充当，而且有不少字的本义是不常用的，如“纳”和“汝”。不然，一方面减少了某些字义混淆的可能性，一方面又会增加另一些字义混淆的可能性。

2. 利用异体字分化职能

异体字是形体冗余现象，利用异体字的不同形体分担不同的职能，使几个形体各司其职。例如：

箸—著—着　“箸”，小篆作“箸”，隶变时，有的将上边的“艸”写作“⺮”，有的将曲笔拉直写作“艹”，造成异体字“箸”和“著”。“箸”的本义是“筷子”，后又记录“显著”义，于是就利用异体字“箸”和“著”来分担这两个职能。在行书中，“著”上边的“艹”可以连写作“⺶”，将其楷化，即为“着”，这样“著”和“着”成为异体字。“著”本来承担了“显著”的意思，又用来表示动作进行状态的虚词“着”，于是就利用异体字分工，由“着”来专门承担虚词义。

亨—享　“亨”和“享”本是一字异体，用法并无区别。本来，祭享、享受的“享”也可写作“亨”，如《易·大有》：“公用亨于天子”。亨通的“亨”也可写作“享”字，如：《周易·乾卦》“元亨利贞”的“亨”，马王堆汉墓帛书和东汉时代的张公神碑都作“享”。大概从唐代，两字的分工才基本固定。

3. 用新造字来分化职能

有些字在有了常用的假借义或专用的引申义后，会另造一个新字来承担它的本义，大量的分化字是这样产生的。而且，由于汉字形声化的趋势日渐明显，这些分化字又大部分是由源字加一

个形声偏旁组成的。

（1）分化本义与借义。例如：

孰—熟　“孰”本义是把食物煮烂，用作虚词，当“谁”讲，它的本义另造“熟”字来记录。

然—燃　“然”的本义已是“燃烧”，借作虚词后，本义另造“燃”字来分担。

莫—暮　“莫”的本义是“晚上”，借作否定词后，本义另造“暮”字来分担。

以上三例是本字被借用后，本义和借义共用一字，后来为了区别，本义反而另造新字分化出去，借字占用了原字。也有另一种情况：本义与借义共用一字，后来为了区别，分化出去的是借义而不是本义，如：

舍—捨　“舍”的本义是“屋舍”，借作“舍弃”义，两词共用一字。以后又造“捨”字，将借义分化出去。

象—像　“象”的本义是长鼻动物的名字，借作“像似”字后，两义共用一字，后来为了区别，另造“像”字将借义分化出去。

（2）分化本义与引申义。例如：

当引申义已与本义的距离较远时，也会产生分化字，或引申义分化出新字，本字仍承担本义；或本义分化出新字，源字反而被引申义占用。

①引申义分化出新字。例如：

坐—座　“座位”义本是“坐立”义的引申，都写“坐”，后来才分化出“座”字承担“座位”义。

齐—剂　“齐”的本义是“平齐”，引申为“协调”义，又引申为“剂量”义，分化出“剂”字。

尉（熨）—慰　“尉（熨）”的本义是将织物熨平，引申而有“安慰”义。分化出“慰”字。

竟—境　“竟”的本义是乐曲的终结，广度加大有一般终了的意义。引申为疆土的边界，分化出“境”字。

②本义分化出新字。例如：

要—腰　“要”的本义是“腰”，引申有“要害”的意思，且“要害”成为“要”的常用义后，于是另造“腰”来承担它的本义。

止—趾　“止”本义是“脚趾”，后来引申义“停止”成为“止”的常用义，于是另造“趾”来承担本义。

由于引申而产生的分化字，源字与分化字之间有同源的关系。

三、同源通用字

具有同源关系的字，即源字和分化字或分化字与分化字，在记词时互相通用，即为同源通用字。

随着人们社会生活的日益丰富，新词便会随着词义的引申而不断派生。起初，这些新词是由源词的字形来承担的，但一字多义终究不利于交际，因此，在词义的推动下，就会不断由源字而分化新字。源字和分化字或分化字之间具有同源关系。当分化字已承担了源字分化出的新义，与源字有了明确的意义分工以后，由于过去长期的习惯，在新词尚未被完全习用的阶段，仍有分化字与源字混用的情况。同一源字分化出的两个以上的新字，也可能在过渡阶段因分化未成熟，尚未成为多数人的习惯，而出现分化字相互混用的情况。这就是同源通用现象。例如：

风—讽　“风”的本义是指自然之风。引申而有“风化”义。《诗·序》：“上以风化下，下以风刺上。”“下以风刺上”即今天所说的“讽谏”。后来又造了一个“讽”字来承担“讽谏”这个意义。但在《汉书》中，“讽谏”义都写作“风”，凡几十见，且注作“读为讽”。

正—政　“正”有“正直”义，引申而有“教化”、“治理”的

意思，后来造“政”字来表示这些引申义。但文献中常有以“正”代“政”的现象。如《左传·文公六年》“弃时政也”，《昭公十五年》“以为大政”，在《汉书》中“政”都写作“正”。

取—娶　“取”的本义是捕获，《说文》：“取，捕取也。”引申指娶妻。《诗经·齐风·南山》：“取妻如之何，必告父母。”后专造“娶”字表此义。《左传·隐公元年》：“郑武公娶于申，曰武姜。”但文献中多有以“取”代“娶”者。如《玉台新咏·古诗为焦仲卿妻作》：“终老不复取。”褚少孙：“巫行视小家女好者，云是当为河伯妇，即娉取。”

说—悦　《论语·学而》：“有朋自远方来，不亦说乎？”这里的“说”当“高兴”讲。“说”是将道理解释开，“悦”是因想开而喜悦，“说”与“悦”同源而经常通用。

同源通用字与同音借用字在形式上相似：第一，它们都不写本字而写它字；第二，本字和通用字以音同或音近为条件。两种现象的不同之处是：同音借用字之间没有意义关系，两者音同音近而义不通；同源通用的双方意义有渊源关系，音同音近而义通。

这是两种本质不同的现象：同音借用是文字使用中的表音趋向的反映，是单纯的文字现象，只是因为本义与借义的形体混同而造成意义混淆；而同源通用是语词分化推动文字分化的过渡阶段源字与分化字界限模糊的表现，它不仅仅是文字现象，同时也反映词汇现象。

第五章
汉字的规范化

第一节　汉字的正字法

正字法，指用社会公认的统一的标准形体来规范文字的书写和使用。每一种文字都有社会公认的统一的标准形体；凡是有文字的国家或民族，一般都有自己的正字法，只是不同体系的文字有着不同的正字法。

拼音文字的字与词基本对应，字形的问题不是很大，正字法的着眼点主要是词，因此又称正词法或正写法。拼音文字的正词法主要解决词的拼写问题，如规定字母表达音位的方法、词的定型化（分写、连写和半连写）、大写规则、移行规则，等等。表意文字的汉字最突出的是形体问题，因此，汉字的正字法主要规定字的正确写法和字体。文字必须确立统一的标准形体，这是文字规范化提出的要求。文字是记录语言的符号系统，必须有全社会统一的标准和规范，否则会影响文字职能的发挥，作为社会交际工具的文字，必须确立形体和书写的统一规范；使用文字也必须符合社会所确立的这个标准，这样才能保证文字的职能得到准确无误的发挥。

一、历史上的正字法

汉字起源早，汉字的正字法也源远流长。文字是社会书面交际的符号系统，字形的不规范势必影响其社会职能的发挥。因此，可以说，从汉字系统产生的那天起，正字的要求就随之而产生。据学者研究，远古的黄帝时期，是原始汉字萌芽的时代；黄帝的史官仓颉，是传说中的进行正字工作的第一人。《荀子·解蔽》中说："好书者众矣，而仓颉独传者，壹也；好稼者众矣，而后稷独传者，壹也；好乐者众矣，而夔独传者，壹也；好义者众矣，而舜独传者，壹也。"

壹，指专一。关于"仓颉造字说"，同后稷对农业、夔对乐等传说一样，应当理解为对相关事物有重要贡献。古代"好书者"很多，惟有仓颉大名独传，说明仓颉对汉字的整理和规范作出了贡献。章太炎在《造字缘起说》中指出：人们造字，"字各异形，则不足以合契。仓颉者，盖始整齐划一，下笔不容增损。由是率尔著之符号，始为约定俗成之契。"章氏的说法是很有道理的。

周宣王时，太史籀整理大篆，编成《史籀篇》，作为教贵族子弟识字的教材，同时也是社会用字的规范。秦代的"书同文"，是历史上第一次由官廷倡导的正字运动。经过春秋战国时期的社会大动荡、大分裂，秦王朝面临的是"言语异声，文字异形"的局面，直接影响了中央统一政令的颁布施行。于是，秦始皇推行"书同文"政策，用秦系文字统一官方应用文字。从两汉到唐代，统治阶级多次立石刻经，选派书法家文化官员用标准的字体书写，作为当时文字的字形规范。科举盛行以后，历代又出版了很多正字的字书，如唐颜元孙的《干禄字书》、张参的《五经文字》，宋张有的《复古编》，清龙启瑞的《字学举隅》，等等。科举考试对字体的要求很严，只要主考官在考卷里发现一个错别字或"破体

字、俗字”，文章写得再好，考生也不会被录取。朝廷对于官吏的文书，在字形规范上要求也非常严格，如果发现奏章上写了错字，轻的打板子，重的还要降职办罪。国家政府对汉字正字的重视和参与，促进了汉字的规范化，促使汉字系统的统一和完善。

在我国历史上，不论哪一个朝代，所维持的正字都是该时代通行的占统治地位的字形，其正字法都是在照顾传统的基础上着眼于当代，着眼于方便运用。

例如，秦朝的“书同文字”，一方面，用秦国文字统一六国文字，“罢其不与秦文合者”；另一方面，传统文字大篆“或颇省改”，以通行的小篆为标准文字。这种政策和原则，顺应了社会对文字的要求。秦在规定书面用标准字体的同时，承认隶书的辅助字体的地位。对官方文字与民间应用文字作出不同的规定，在当时是有进步意义的文字政策。唐代仍然奉行这种政策，如颜元孙《干禄字书》，把汉字分为“俗、通、正”三体，分别规定其不同的应用范围。

汉承秦制，最初以小篆为官方文字，秦代推行小篆的范本《仓颉篇》等，仍为汉初“闾里书师”教学童的课本。随着民间隶书的通行和完善，汉代扩大了隶书的使用范围，并在东汉最终扬弃了小篆，使在民间通行已久的、更便于书写的隶书成为官方标准字体。东汉蔡邕发起并奉敕刊刻的《熹平石经》，代表的就是隶书的标准体。

汉末六朝，社会动荡，政权更迭频繁，文字的规范力度减弱；同时，由于书面交际范围的扩大和文字自身体系的发展，讲究蚕头燕尾、点画波折的隶书已经不适应书写简便快捷的要求，书写更快捷、更方便的楷书在民间兴起并趋于成熟。南朝梁五帝、简文帝父子政治上算不上有为之君，但还能在社会短暂的稳定时期关注文字工作，命顾野王编纂了第一部楷书字典《玉篇》，从而奠定了楷书的正统地位。唐代，楷书正式成为官方文字，颜师古以

来的“字样”之学，以正楷刊石的《开成石经》，都起了官方正字法的作用。

楷书基本上适应了汉字在书写上求简便、在目治上求区别的要求，因此，唐代的正字标准，成为以后宋元明清各代的正字基础。宋元以来千有余年，都采取了尊重传统、维护汉字的稳定和统一的正字原则。至清代，皇帝钦敕编纂《康熙字典》，成为汉字正字的总结之作，为汉字的定形、定音、定义作出了贡献。

综上所述，我国在正字法方面有着悠久的历史和优良的传统，这也是汉字系统历经几千年，构形系统日益完善、全社会用字保持统一的重要原因。古人的正字历史为我们留下了可供借鉴的经验教训。

二、新中国成立后的正字工作

尽管历代统治者重视，汉字学者作了不懈的努力，然而在封建时代，由于种种原因，汉字规范化的程度依然很低，字形存在着严重的分歧，具体表现为繁简并用、异体并存、印刷字形不统一。建国以后，党和政府亲自领导了汉字的整理和简化工作，建立并全力推行了科学适用的正字法，实现了前所未有的长时间的广范围内的书面用字的规范和统一。

建国后的正字工作主要有四个方面的内容：

（一）简化汉字，确立简化字的正字标准

确立文字的形体标准，是一个历史的范畴，不是永恒不变的。因为文字总是在不断演变和不断规范的矛盾中发展的。文字的社会交际职能要求文字形体固定，要求文字的统一和标准化，用正字法来限制它的变异和不规范。可是，文字在人们的使用过程中，又不可避免地会突破正字标准的约束，在一定程度上使汉字体系

发生变化。有些变化是合理的，符合汉字发展趋势的，有些变化则是违反汉字规律、破坏汉字系统的。在这种情况下，社会又不得不重新确立正字法的标准。因此，不同的历史阶段具有不同内容的、不同标准的正字法。新中国成立后所确立的正字标准，一部分是沿用过去的通行字，另一部分是对笔画过多的字进行简化。

1. 简化汉字的理论依据

汉字是因义构形的表意文字，其形体直接携带着意义信息。汉字既是表意文字，便具有表意文字必然具有的特点。首先，当人们根据意义为词或语素指定视觉符号时，总是与造字者的具象思维相联系，抽象性不可能很强，而且，越是要求符形的表意能度高，携带的信息量大，就越会增加符形的难度。文字符号必须保持符形的区别度，而意义之间的差异纷繁而模糊，一个一个去区别，势必使符形更加繁化。所以，表意文字个体字符丰满而繁难，在它的早期往往不可避免。其次，表意文字的性能是形义统一的，而词义是语言中演变最快、最活泼的要素。意义的增多和改变时时会推动字符的增多和符形外观的改变。所以，一种文字如果坚持表意性能，它的字符数量必会越来越多。但是，符形的繁化阶段在文字的使用范围日益扩大以后，便必然要结束。这是因为，文字职能的发挥，是由书写和认读这两个不可缺少的环节合成的。就认读而言，人们希望符号形象丰满、易于识别；就书写而言，又总是希望符号简单固定，便于书写。追求形象丰满，往往会增加符形的繁难，不符合书写的要求。特别是当字符数目激增、数量超过了人们的记忆负荷之后，符形的繁难就更难被广大的使用者接受。为了克服数量激增和造型繁难这两个尖锐问题，作为表意文字的汉字必然要遵循以下两个规律发展：

第一，在易写与易识的相互矛盾中，个体字符要不断调节造型，使其简繁适度，达到优化的标准。也就是说，在保证必要的表意性和区别度的前提下，要最大限度地减少构件和笔画数。

第二，在新词新义大量增加、字符数量不得不随之增加的情况下，不断地规整基础构件与基本构字方式，使文字的总体构形系统日趋完善与简化。

以上两个规律，是在自然发展与人为规范的交替中实现的。

汉字在历史上经过多次的人工规范，例如李斯对小篆的整理、东汉对隶书的整理、南北朝、唐代对楷书的整理，其中最彻底、最系统的一次是东汉许慎对前代小篆体系的人工规范。每一次规范，对前代文字来说，都是一次系统的简化。小篆是对大篆的简化，隶书又是对小篆的简化。自《说文》对汉字的结构体系进行规范之后，汉字的结构体系基本定型，汉字的系统简化减缓，但是，个体字符的简化从未停止过。从商朝开始，人们为了使汉字便于应用，一直在简化汉字的形体。甲骨文、金文中有简体字，隶书、楷书中简体字更多。唐宋以后，由于印刷术的普遍使用，民间书坊的增多，简体字的传播更为迅速。近代，太平天国将简体字定为通用字体，连玉玺都使用简体字。这是在中国历史上简体字第一次被官方确定为社会通行文字。

2. 简化汉字的历史进程

进入 20 世纪以来，社会上出现了更多的简体字。同时，许多以普及文化、开发民智为救国手段的知识分子，开始收集研究和提倡简体字，并提出各种简化方案。较有影响的有：1921 年陆费逵的《整理汉字意见》、钱玄同的《减省现行汉字的笔画案》。陆氏提出了整理汉字的办法：限定通俗用字，减少汉字笔画；钱氏则主张把简体字广泛应用于一切正规的书面语。此后，出现了一批收集整理简体字的专书，如钱玄同、李家瑞的《宋元以来俗字谱》（1930），杜定友的《简字标准字表》（1934），钱玄同等的《简体字谱》（1935），容庚的《简体字典》（1936），等等。在文化界进步人士蔡元培、邵力子、郭沫若、郑振铎等人的积极倡导下，简体字成为一些正式报刊的用字。在简体字运动的推动下，1935

年8月，国民政府教育部正式公布了《第一批简体字表》。这是历史上由中央政府公布的第一个简体字表。但是，由于当时简化汉字的条件还不成熟，这个字表又被通令收回。这个字表虽然为时很短，但对社会上应用、流行简体字仍然产生了很大的作用。

综上所述，建国后开展的汉字简化运动具有广泛的社会基础，是适应汉字系统发展的趋势，顺应历史潮流的正确举措。

新中国成立以后，面临着百废待兴的局面，党和政府立即着手组织专家对汉字进行有计划的简化。这次简化工作的总原则是："约定俗成，稳步前进"。约定俗成，是指在社会习惯的基础上，因势利导，尽量采用已经流行的简体字。具体作法是：在过去汉字简化的基础上进行简化，首先整理、研究和肯定在群众中长期而广泛流行的、已经社会化了的简体字，对此只做必要的修改和补充；简体字的选择以最常用的为限，不是对每一个繁难的字都简化。稳步前进，是指对需要简化的字分批进行，不一次解决，也不一次推行。这样做既符合社会需要，又有群众基础；既便利初学者的学习，又照顾已识字人的习惯，从而保证简化字得以顺利推行。

3. 简化汉字的具体操作方法

（1）保留轮廓，减省笔画。即省去原字中繁难的部分，只留下轮廓部分。如：

蓋——盖　傘——伞　樹——树

齒——齿　龜——龟　齊——齐

（2）以点代面，减省部件。即选取其中的一个部件来代替原字，省去其他部件。这也可称作"以部分代全体"。如：

聲——声　開——开　糶——粜

醫——医　氣——气　遲——迟

（3）声符义符，以简换繁。即用构形简单的声符或义符替换原来的构形繁复的声符或义符。

替换义符的，如：

願——愿　　鹼——硷　　骯——肮　　貓——猫

替换声符的，如：

殲——歼　　劇——剧　　擁——拥

億——亿　　燈——灯　　種——种

（4）就地取材，草书楷化。即将已经社会化的草书字的笔形楷化，作为规范字。如：

書——书　　貝——贝　　車——车

專——专　　東——东　　會——会

（5）偏旁改造，符号替换。即用一个笔画简单的符号来替代一个或几个笔画繁难的偏旁。这个符号只起替代作用，没有表意或示音功能。如：

僅——仅　　漢——汉　　勸——劝

區——区　　學——学　　興——兴

（6）全面改造，另造新字。有些字不便简化，亦不便以同音字替代，就另造新字。这些新造的字有两种情况：一是新造字为会意字和形声字。如：

塵——尘　　寶——宝　　體——体

竈——灶　　衆——众　　（以上会意字）

驚——惊　　態——态　　郵——邮

審——审　　護——护　　（以上形声字）

二是新造字为纯粹的符号。如：

義——义　　頭——头　　萬——万

幣——币　　辦——办　　棗——枣

（7）古为今用，弃今用古。即抛弃现用字形，采用古字来取代今字。如：

雲——云　　纔——才　　從——从

電——电　　鬍——胡　　鬚——须

（8）同音借用，鹊巢鸠占。即借用结构简单的音同或音近的字，来替代结构繁复的字。采用同音替代简化法的前提是：借用字与被借用的字，其中有一个必须是罕用的僻字，借用以后不至于发生意义的混淆。这种方法如果选字得当，既简化了笔画，又精简了字数。如：

醜——丑　　裹——里　　臺——台

韆——千　　鼕——冬　　齣——出

（9）举一反三，偏旁类推。即用简化偏旁或可作偏旁的简化字，来替代合体字中的同形构件，以类推出一批简化字。

上述简化的原则和方法不是凭空想出来的，而是在研究、总结前人简化汉字的经验的基础上提出来的。虽然简化过程中有一些具体操作方法还存在一定的不足之处，但其成就是主要的，大方向是准确的。

（二）整理异体字，确定正字标准

异体字是指结构不同而音义全同，具有相同的记词功能的一组字。汉字自古以来存在着异体。异体字的存在造成了一字多形，增加了学习和使用的负担，影响了汉字职能的有效发挥。随着汉字使用范围的日益广泛，要求不断提高规范化的程度。整理异体字，有利于减少字数，逐步消除字形混乱现象。因此，全面整理异体字便是汉字正字法的重要内容。建国初期，在简化汉字的同时，国家有关部门立即着手整理异体字。

这次整理异体字主要本着两个原则，就是从俗和从简相结合，照顾书写方便。

（1）“从俗”是指选择应用较广的字，废除较生僻的字。例如：

选用的规范字：村　乃　奔　吝　冰　简　拖

废除的异体字：邨　廼　犇　悋　氷　筩　拕

（2）“从简”，就是保留笔画较少的，废除笔画较多的。例如：

选用的规范字：采　笋　你　志　晢　挂
废除的异体字：採　筍　妳　誌　晣　掛

（3）如果从俗和从简不能兼顾，或繁简相差无几，则以从俗为主。例如：

选用的规范字：耻　考　游　够　同　阔
废除的异体字：恥　攷　遊　夠　仝　濶

（4）有左右和上下结构的异体字，为了便于书写，一般选用左右结构的字作为规范字。例如：

选用的规范字：峰　群　棋　晰　略　鹅　峨
废除的异体字：峯　羣　棊　晳　畧　鵞　峩

（5）只有少数的几组，因为群众习用，仍选用上下结构的字作为规范字。例如：

选用的规范字：拿　蟹　幕　岸
废除的异体字：拏　蠏　幙　岍

"从俗"符合文字的社会性，"从简"符合文字发展的主要趋势。因此，这两个原则基本上是正确的。

（三）整理印刷体字形，确定印刷字形规范

整理印刷体字形，主要是对异写字进行整理。异写字是指构字部件相同，而写法不同的一组字，如"刻"即有"𠛮、㓩、刻"等不同写法。在字形整理以前，汉字的楷书印刷体中，存在着许多字形不统一的现象，即同一个字，在不同的出版物中，写法多有出入。有些字在不同的旧字典中，形体也不一样。例如，青——靑、花——花。这些异形，主要是由于是书写造成的异体，因而是异写字。异写字的存在，使得部首、笔画、四角号码等汉字排检法都遇到许多不易解决的难题。印刷体的混乱现象同样也影响到手写体，给人们学习和使用汉字增加了负担。同时，汉字信息的计算机处理技术也需要汉字字形的标准化。为了适应出版业

的发展，减轻社会学习和用字的负担，使汉字更好地适应信息处理科学技术的进步，整理印刷体字形，实现汉字字形的标准化已是刻不容缓。为此，当时的中国文字改革委员会和文化部联合对印刷体字形进行了整理。这次字形整理的总原则，仍然是“从简从俗，便于学习和使用”，具体操作原则，主要有以下八点：

（1）笔画繁简不同的，选用笔画简省的。如（规范字形——淘汰字）：

吕——吕　黄——黃　巨——巨　宽——寬

（2）笔画有连、断不同的，选用笔画连接的。如：

花——花　敢——敢　片——片　免——免

（3）笔画有中断与延伸不同的，选用笔画延伸的。如：

角——角　灰——灰　蝇——蝇　茇——茇

（4）笔画有短长之别者，选用笔画缩短的。如：

丑——丑　那——那　亏——亏　周——周

（5）起笔笔画有撇横之别的，选用横笔的。如：

丰——丰　刊——刊　邦——邦　忝——忝

（6）起笔为点的，一律写侧点。如：

主——主　永——永　户——户　令　令

（7）笔画有直笔和折笔之别的，选用直笔。如：

吴——吳　直——直　缜——缜　普——普

（8）笔画有八、ⅴ之别的，选用“ⅴ”的。如：

关——关　平——平　半——半　肖——肖

从以上操作原则可见，印刷字体的整理原则，在保证“从俗从简”的前提下，尽可能地考虑到了书写的方便，如直笔、侧点、“ⅴ”笔形的选择，弃“角”写“角”的规定，从而减少了写错字的几率。

简化汉字、整理异体字、整理印刷体字形，是我国正字工作的主要内容。体现这些工作成就的，是下列代表着国家文字规范

的标准和依据的几个字表（见第二节）。

三、其他正字工作

除了上述几项重要的字形整理工作之外，我国还对地名用字、计量用字进行了字形规范。

1. 改换生僻地名用字

我国幅员广阔，许多地名用字地域性很强。这些字在当地属常用字，但在全国范围内是生僻的，其中还有一些字笔画繁复，书写不便。这种状况，给人们的社会交往和信息处理造成了障碍。同时，这些字的存在，给计算机的汉字输入造成了困难，也在邮电传递上发生错误。因此，从 1951 年 12 月 19 日发出更改地名用字的指示开始，到 1964 年止，国务院批准 35 个县级地名中的生僻字改用为常用字。这些生僻的地名用字，有的难认或容易读错字音，例如，“雩都县”改为“于都县”，“新淦县”改为“新干县”，“盩厔县”改为“周至县”；有的笔画太多，例如，“鳛水县”改为“习水县”，“醴泉县”改为“礼泉县”，等等。

1987 年 3 月 27 日，国家语委、中国地名委员会、铁道部、交通部、国家海洋局、国家测绘局联合颁发了《关于地名用字的若干规定》的通知，对地名用字规范作了新的规定，要求“地名的汉字字形，以 1965 年文化部和中国文字改革委员会联合发布的《印刷通用汉字字形表》为准”。

2. 整理计量单位名称用字

1959 年，国务院发布了《关于我国计量制度的命令》，纠正了一种术语表示两种计量单位的状况，废除了一些特殊汉字，如“瓩、粴、噉”等，对改变名称计量单位用字的混乱状况起到了积极作用。但是，计量单位名称用字仍有混乱现象。如在书面语中，“盎司”与“温斯”并存，“两”存在着“liǎng、yīngliǎng”两种

读音；还有一些复音字，如“吋”读 yīngcùn”，等等。

1977 年 7 月 20 日，国家文字改革委员会和国家标准计量局颁布了《关于部分计量单位名称统一用字的通知》。《通知》要求，所有出版物、打印文件、设计图表、商品包装，以及广播等，均应采用“附表”中选用的计量单位译名用字，停止使用其他译名用字。《通知》所附的《部分计量单位名称统一用字表》，废除了一字双音即复音字如“哩、呎、吋”，分别改用“英里、英尺、英寸”。

第二节　国家的正字标准

国家的正字标准，是指政府有关部门所规定的社会通用的字形标准，其具体内容是以每项正字工作后所公布的字表来体现的。

一、《简化字总表》

1964 年 5 月中国文字改革委员会编辑出版。该表是解放后汉字简化工作的总结和成果的集中体现。全表由三个字表组成：第一表收 352 个不作简化偏旁用的简化字；第二表是 132 个可作简化偏旁用的简化字和 14 个简化偏旁；第三表是应用第二表所列简化字和简化偏旁类推简化得出来的 1 754 个简化字。三个表总计为 2 738 字。其中“签、须”两字重现，实为 2 736 字。《简化字总表》是共和国政府领导的、由专家学者参与的第一次汉字简化工作的最后总结，被确定为社会通用汉字的正字标准。要正确使用这些字，必须认真学习《总表》，特别是第一、第二表。具体应注意以下问题：

（1）第一表的字是不能作为简化偏旁用的，因此不能类推简

化其他繁体字。例如：

"習"简化作"习"，而"褶"的右偏旁不能类推简化；

"兒"简化作"儿"，而"倪、霓、猊、锐、鲵"等字的右偏旁都不能简化作"儿"；

"幹"简化作"干"，而"擀"却不能简化作"扞"，等等。

第一表的字一般都是个别简化的，因此不能用第二表的字将它们分别类推简化。例如："節"简化为"节"，但"癤"不能类推简化为"疖"，而是简化为"疖"，等等。

对不能类推简化的字，《总表》中都作了特别说明，在学习时一定要注意。

（2）第二表的字都可以用来类推简化其他繁体字。但是，决不能用来类推简化第一表中的繁体字。这些字都已经作了个别简化，需要逐个记忆。

（3）第三表的简化字是用第二表的简化字和简化偏旁类推出来的。没有收入《总表》的字，可以按照应用第二表中的简化字和简化偏旁类推简化。

由上可见，第一、第二表中的简化字需要逐个记忆；第三表的字只要会类推即可，不需硬记。

70年代，我国进行了第二次汉字简化。1977年12月20日，《第二次汉字简化方案（草案）》公布，内收整体简化字462个，连同偏旁类推简化字，共853个。由于这个"草案"单纯从减少笔画出发，违反了汉字的结构规律，忽略了文字的社会性和稳定性，忽略了约定俗成的原则，因此，"草案"既缺乏科学性，又没有社会基础。所以，"草案"公布后，便收到了来自各方面的不同意见，反对的呼声很高，不久便在中小学课本中停止使用。1986年，经国务院批准正式废止。

为了纠正80年代末期以来重新出现的社会用字混乱的现象，便于人们使用规范的简化字，经国务院批准，国家语委于1986年

10月10日重新发表了《简化字总表》。根据几十年来简化字在社会上的使用情况，新版的《简化字总表》对原字表中的个别字作了调整。调整的具体内容是：

第一，“叠、覆、像、啰”不再作为“迭、复、象、罗”的繁体字处理。因此，在第一表中删去了“迭（叠）”、“象（像）”；在“复”字头下删去了繁体字“覆”；在第二表“罗”字头下删去了繁体字“啰”，“啰”依简化偏旁类推简化为“啰”。

第二，“瞭”字读“liǎo”（了解）时，仍简化作“了”；读“liào”（瞭望）时，不简化。

国家语委在《关于重新发表〈简化字总表〉的说明》中强调，汉字的形体在一个时期内应当保持稳定，以利应用，并“要求社会用字以《简化字总表》为标准，凡是在《简化字总表》中已经被简化了的繁体字，应该用简化字而不用繁体字；凡是不符合《简化字总表》规定的简化字，都是不规范的简化字，应当停止使用”。

二、《第一批异体字整理表》

异体字的整理工作既是整理字形，又是减少字数。这项工作几乎是同汉字简化同期进行的。整理的结果，是1955年12月，文化部和中国文字改革委员会联合公布的《第一批异体字整理表》（简称《异体字表》）。

《异体字表》收异体字810组，共1 865个字，每组只留一个正体，其余作为异体淘汰，共淘汰1 055个异体字。该表自1956年2月1日起在全国正式实施，全国出版的报纸、杂志、图书，一律停止使用表中括号内的异体字；机关、团体、企事业单位用的打字机盘中的异体字也逐步改正。只有四种情况可以作为例外：①翻译古书需用原文原字的；②一般图书已经制成版的或分册尚

未出齐的，可以重排再版时改正；③商店原有牌号不受限制；④用作姓氏的，在报刊图书中可以保留原字。

根据汉字使用的实际需要，此后对《第一批异体字整理表》的内容又作了局部调整。调整的具体内容是：

（1）根据1986年10月10日重新发表的《简化字总表》的说明，确认该表收入的下列11个类推简化字为规范正字，不再作为淘汰的异体字：

䜣　䜩　晔　詟　䲡　䌷　刬　鲙　诓　雠

（2）根据1988年3月25日国家语言文字工作委员会与国家新闻出版署《关于发布〈现代汉语通用字表〉的联合通知》中的规定，确认《印刷通用汉字字形表》收入的下列15个字为规范字，收入《现代汉语通用字表》，不再作为淘汰的异体字：

翦　邱　於　澹　骼　彷　菰　溷　徵　薰　黏

桉　愣　晖　凋

经过上述调整，《第一批异体字整理表》实际淘汰异体字1 427个。《异体字表》的颁布和实施，有效地减少了出版物上的异体字，精简了社会通用字的数量，给学习和使用带来了方便，受到了社会各界特别是出版、教育界的欢迎。

三、《印刷通用汉字字形表》

该表是中国文字改革委员会和文化部联合对印刷体字形进行整理的结果，于1965年1月公布（简称《字形表》）。

该表收字6 196个，遵循从简从俗、便于学习和使用的原则，提供了通用汉字印刷字体（宋体）的标准字形，明确规定了表内字的笔画数目、笔画形状、笔画顺序和构件部位。它既是印刷字体的标准，也为手写体确立了字形规范，同时也是写字教学的标准。

《字形表》的发布，是中华人民共和国建立后汉字整理工作的重要成果，对统一印刷字形，促进社会用字规范化和方便中文信息处理都具有重要意义。

四、《现代汉语通用字表》

这是全面体现新中国成立以来正字工作成果的字表。该表于1988 年 3 月 25 日由国家语委和国家新闻出版总署联合发布，共收字 7000 个，是国家公布的最新的规范字表。它规定了每个字的规范字形，包括笔画数、笔顺和笔画部件的组合结构。这个表体现了《第一批异体字整理表》、《简化字总表》、《印刷通用汉字字形表》等重要字表的基本内容，吸收了各个字形规范化字表的精华，并作了某些合理的调整。因此，此表可作为国家正字工作的主要标准。语文出版社出版了《现代汉语通用字表》一书；其他几个字表可参看该社出版的《语言文字规范化手册》一书。

第三节 书写规范

所谓“书写规范”，是指在书写汉字时，其字形符合现行汉字“正字法”的规定。

书写规范，是汉字书写的基本要求，是出于实用目的而非艺术的目的提出来的。文字是为扩大语言的时空局限，扩大语言的交际范围而诞生的，是记录语言的符号系统。在印刷术发明以前，文字主要是通过人工书写来实现其书面交际的功能的。汉字也是如此。印刷术的发明，特别是 20 世纪末以来我国计算机的普及并普遍用于汉字信息处理，使汉字的人工书写量大大降低。但是，人工书写汉字在今后一个相当长的时期内将仍然是非常重要的书面

交际形式；特别是在语文教育工作中，更难以离开汉字的人工书写。如果汉字的书写错讹、混乱，必然妨碍汉语的交际功能，影响语言文字的规范化。因此，汉字的书写规范一直是汉字正字法的重要内容。

重视汉字的书写规范，是中华民族的优秀传统。早在周秦时期，就有了所谓“书学”，即学习写字。学习写字的重要内容就是书写规范。汉代，书写规范成为统治者对知识分子的基本要求。《汉书·艺文志》：“吏民上书，字或不正，辄举劾。”《说文·叙》：“（尉律），书或不正，辄举劾之。”“书”就是字。在这种严格的要求下，知识分子把写字规范看得非常重要，写了错字，被当作很严重的事。据《史记·万石君传》：“（石）建为郎中令，书奏事。事下，建读之，误写‘马’字。‘马尾当五，今乃四，不足一；上谴，危矣！’甚惶恐。”一个堂堂的国家大臣，只不过写了一个错字，便担心“上谴”，因此“甚惶恐”，除了石建本身的过于谨慎之外，也说明了当时对汉字书写规范要求的严格。

在当今，文字已被绝大多数人掌握，汉字书写已经成为一种普遍的技能，因此，汉字的书写规范将显得更为重要。

书写规范主要包括以下三方面的内容：

一、字形规范

字形规范是书写规范的总体要求，是指写出来的字形要符合现行汉字的字形规范。具体要求是：

1. 不写繁体字

不写繁体字是指在一般场合下，写规范的简化字，特殊场合，如古籍整理、书法创作、海外或外籍人士题字等除外。现行汉字的正字是规范的简化字，以1986年国家语言文字工作委员会等单位颁布的《简化字总表》为标准。1977年公布的第二批简化字已

经被废除，在公共场合应停止使用，同时，也不要乱造简化字。近来一些人热衷于写繁体字，造成汉字书写的混乱。这是因为：第一，简化字已经推行了 40 年，凡 50 岁以下的人，受的都是简化字教育，写繁体字是一种扬短避长的自我勉强行为，当然不如写了几十年的简化字熟练和顺手，写错别字的现象便随之大量增加；第二，简化字和繁体字之间并不是简单的一对一的关系，二者也决不是两种对立的文字体系，简化字并没有对所有汉字都进行了简化。有些人对此常识并不了解，随意进行繁简转换，出现了许多有辱斯文的笑话。例如，烟台一家糖果厂曾经生产过一种金色锡箔纸包装的元宝形巧克力，包装纸外面用繁体字写着“恭喜髮財”。他们不知道“发”字对应的是两个字，一个是“射箭、出發、發出、發财”的“發”，一个是“头髮”的“髮”。生产者用“髮”代“發”，闹出笑话。“慾”和“欲”互相转换也常常出错。这两个字在繁体字系统中根本就是两个字，前者表示“欲望、想要”，是一种心理活动，在词义色彩上属中性；后者则表示“贪欲、欲望”，是贬义词，二者在古文中决不相混。在简化字中，才以“欲”代“慾”，二者合流。繁简字的错误转换，甚至在高等学府也时有出现。例如，某大学“中文系”的牌子上赫然写着“中文係”，系与係字也犯了繁简转换错误。这种情况，让熟悉简化字的“海内人士”不知所云，让使用繁体字的“海外人士”莫名其妙。

2. 不写被淘汰的异体字

凡是《第一批异体字整理表》中淘汰的异体字，都属于不规范汉字，没有特殊需要不应在社会书面交际中使用。

3. 不写旧字形

新旧字形的界限以 1964 年公布的《汉字印刷通用汉字字形表》为准。本表所收字形为新字形，此前的与此表字形不符的为旧字形。这个字表不仅是印刷字形的规范，同时也是手写体的字

形规范。每个汉字的使用者都应该写规范的新字形，不写旧字形。

4. 杜绝写错别字

错别字是错字和别字的总称。

(1) 错字，是指写的不成字的“字”，指随意增简字的笔画、写错字的偏旁部件、写错字形结构，从而造成错字。写错字，大都是由于对汉字的笔画安排和结构形式辨别不清、记忆不准造成的。例如，把“武”字的右偏旁记成“戈”，把“冠”的上面的部件“秃宝盖”记成“宝盖”，把“创”的声符“仓”记成“仑”。另外，《简化字总表》所不载的简笔字都属于错字。

(2) 别字，是指不合规范的同音替代字，或叫白字。也就是本应写甲字，却写成了与甲字的意义毫不相干的同音的乙字。例如：

刻苦——克苦	自暴自弃——自报自弃
诡辩——鬼辩	不卑不亢——不悲不亢
贡献——供献	英雄辈出——英雄倍出
布置——部置	称心如意——趁心如意

有些字不但音近，形也相近，就更易写错，例如：

证券——证卷	病入膏肓——病入膏盲
姿态——恣态	别出心裁——别出心栽
摧毁——催毁	惩前毖后——惩前毙后
提纲——题纲	痴心妄想——痴心忘想
妨碍——防碍	焕然一新——换然一新
刚愎自用——刚复自用	明辨是非——明辩是非

错别字的存在，妨碍人们准确表达思想，影响书面交际，造成对书面意义的误解。误解比不解危害更大。汉字是表意文字，是以形别义的；形体的不同，代表着不同的意义。因此，写了错别字，往往会因一字之差，意义全变。就产生的原因来说，错字与别字不同，错字来源于书写的失误，别字来源于选用字的失误，但

从记录语言的功能上来说，二者都违背表达的原意。由于错别字的存在是人们进行正常的书面交际的障碍，因此，必须避免写错别字，按照规范的正字和国家规定的标准字形书写和使用汉字。

二、如何避免写错字

避免写错字要从两个环节入手：一是构形要素；二是书写要素。构形要素主要指部件，书写要素主要指笔画包括笔形、笔顺和笔画。

1. 构形要素的规范

汉字是由数量有限的部件按照一定的结构模式构成的。从构形层次上来说，部件规范和结构形式的正确是两个不可忽视的环节。

部件是汉字的构形元素，许多错字都是由于部件的书写错误造成的。特别是因部件形体相近而致错。汉字中有许多形近部件，例如：

位——往、这——建、福——裕等，这三组字的左偏旁形体相近，是形近部件；冠——寇的上部偏旁是形近部件；广与厂、日与曰是形近部件。

其中最容易混淆的是“示字旁”和“衣补旁”、“宝盖头”和“秃宝盖。”如：

“冠”字上面常被加点，“裕”字常被写成“示字旁”，“礼”却常被写成“衣补旁”，“延”字的左偏旁上面常被加点，等等。

要防止将部件安错，最有效的方法是正确理解汉字的构意和部件的构字功能。例如：

(1) “示旁”与“衣补旁”。示，本义是祭祀牌位，因而有“祭祀”与“神灵”的意思，凡从“示”之字皆与祭祀、与“神灵”有关：社，土地神；祟，凶兆出；祠，祭神之地；祝、祷，同

神说话；故字皆从示。“衣”，作偏旁皆用其本义。凡与“衣物”有关之字皆从衣不从示。例如，初，本义是“裁衣之始”，故从“衣”：裕，指衣食富足；衬、亵，指贴身的内衣；袜，指“足衣”；襟、袖，皆为衣裳的某部分；故字皆从衣而不从示。

（2）“宝盖头”和“秃宝盖”。“宝盖头”由房屋之形变来，故凡从“宝盖头”者，皆属建筑物或与建筑物有关的字，如“家、宅、宇”等。“富”，屋里有粮仓；字，本义指“产子”，产子必在屋里；宝，家中有玉有贝；宿，止宿之处；宰，“罪人在屋下执事者”（见《说文》），即屋中的奴隶。“秃宝盖”，由蒙头布巾的形象演变而来，本义为覆盖，所从之字皆与帽子、覆盖义有关。如蒙头的“蒙”（古字没有草字头），以及冕、冒上部所从的部件是古“帽”字，即从“秃宝盖”得义。因此，冠，是一种帽子，上面决不会有点；寇，本义是入“室”抢劫，上面的点决不能丢。

（3）“单立人”与“双立人”。二者虽皆称“人”，却意义迥异。与“行走”有关的字皆从“双立人”，如“行、往、径（指人走的路）、循（顺着走）、御（驾驶马车）”，等等。与“人”有关的字皆从“单立人”，如“仇、仙、侮、倡、优、傀、儡”等等。

（4）广与厂。广，本义是房屋之形，从广之字皆与房屋有关。如：府（藏文书之处）、厨、庐、庑、廊、废（房屋颓败）、庖，等等。厂，是山崖的“崖”的古字，凡从“厂”之字皆与山崖、山石有关。如：厝（磨刀石）、崖、岸，等等。崖、岸上面的“山”是后加上去的。

部件位置的规范也很重要。汉字的结构形式有左右、上下、内外三种主要的结构形式和其他几种次要的结构形式。在绝大多数汉字中，部件的位置是固定的。但是，部件摆放的位置，作为一种区别要素，在汉字体系中并不严密。主要有三种情况：①部件摆放位置的不同，构成异体字。在有些汉字中，部件的位置没有区别意义的作用。例如：晰、鹅、群、峰、蛾都有左右、上下两

种结构形式，构成意义全同的异体字。②部件位置的改换可构成错字。有一些字，其部件的结构形式是固定的，不能随意改换；改换了结构形式，就构成了错字。例如，邻、融、鳞，是左右结构，把它们的部件左右调换一下，就会变得不成字；遵，把“寸”拿到最下面，也会成为错字。③部件的摆放位置不同，构成记词功能不同的字。在多数情况下，部件摆放的位置则构成区别意义的手段。即，部件的结构样式不同，就构成记词功能不同的两个字。例如：

杏——呆　杲——杳　旮——旭　旮——旯　晰——晳

由于部件的位置区别意义的作用并不严格，因此给汉字的规范书写带来了一定的困难。在记忆、书写汉字时，一定要注意部件的摆放位置，防止写错字。

2. 书写要素的规范

汉字的书写要素是笔画，笔画有三个要素：笔形、笔数和笔顺。

（1）汉字的笔形。汉字是由不同笔形的笔画一笔一画写出来的。因此，了解汉字的笔画首先要了解汉字的笔形。

现在的规范字体是楷书。前面说过楷书的基本笔形共有六种，即点、横、竖、撇、捺、折（丶一丨丿㇏乙），其他笔画形式都是这六种笔形的变体。有人把捺归入点，因而又有五种之说。又有人把挑（又称提）和竖钩（亅）分出来，因此又有八种之说。其实，挑是横的变体，钩是竖的变体，可以不必分出。

楷书的基本笔形在不同的字里和不同的部位上又有许多变化。这些变化了的笔形，又叫“发展笔形”，与“基本笔形”相对。掌握了基本笔形及其写法，发展笔形便很容易掌握。

每个规范汉字的笔形是确定的，不能改变，否则就会出现错别字。例如，“天”的末笔是捺，写成折钩就成了“无”；“元”字最后一笔是折钩，写成一个斜点就与“六”分不清了。相同形状

的笔画在每个字中的位置是固定的，位置的改换，就成了不同的字，如主——玉、庄——压，点的位置不同，构成不同的字；末——未、士——土，长横短横的不同位置，构成不同的字。可见，汉字笔画与书写规范的关系是相当密切的。

（2）汉字的笔数。每个规范汉字的笔画是有定数的，书写时一定要注意不可随意增减笔画，否则就会产生错字。例如："庆"与"驮"中的"大"字上加一点，便成为错字；"伐"字少写一撇，就成了"代"字。

（3）汉字的笔顺。由于绝大多数汉字是由多笔画构成的，所以书写时就存在一个下笔的顺序问题。这个下笔的先后的规则，就叫笔顺。掌握笔顺规则，目的是为了提高书写速度，避免写错字，并使汉字写得均衡漂亮。

笔顺的规则大致有八条：

①先上后下。如：立、宝、孳、旦

②先左后右。如：林、愧、陈、以

③先外后内。如：用、匡、闷、匀

④先中间后两边。如：办、小、水、承

⑤先外后内再封门。如：圆、回、田、目

⑥先横后竖。如：十、干、井、丰

⑦先撇后捺。如：八、大、入、人

⑧先横后撇。如：厂、左、右、万

汉字的形体多样，因此有些字的笔顺也很特殊，可能有不符合上述规则的情况。在把字写得均衡规范的前提下，可以灵活变通。遵照笔画规则写字，可以把字写得整齐端正。例如"匡"字，先写好外框，内部结构就容易安排了。"丰"字，先把主体位置确定，再从中间穿过，便于把字写得均衡美观。

有些书法家出于某种审美的需要，可以不按笔顺规则书写。但对于一般的汉字书写行为，应尽量遵守笔顺规则。当然，笔顺规

则不是绝对的，在书写一些笔画复杂、特殊的汉字时，可以不必强求一致，只要写着顺手，写出来字形正确、端正就可以。但是，总的运笔规则却是不能违背的。违背运笔规则，写“倒笔画”，就不可能把字写快写好。

第六章
汉字与汉字教学

第一节　我国汉字教学的历史传统

一、汉字教学的起源

汉字是世界上最古老的文字之一，在汉字被应用于政治生活和社会交际的同时，汉字教学就已经开始了。根据考古文献研究，在汉字系统形成的殷商时期，习字学书就已经成为当时教育的重要内容。郭沫若的《殷墟粹编》1 465～1 479 片，收有刻写六甲表的前 12 行，他在《序言》中指出，这是商人教子弟刻写文字的记录。张政烺通过对各代文物的比较考证，肯定了郭说。他指出："我们知道殷墟卜骨上常见学刻六甲表的前几行，居延汉简上常见学写《仓颉篇》、《急就篇》的前几句，汉砖上也常见学写《春秋公羊传》的首章，都是学书者所作，事情或许相同。"① 这说明，商代确有教学习字的"书学"，只是当时的文字主要掌握在巫史手里，汉字教学因而也只是极少数人的内部行为而已。

西周时期，汉字教学已经被正式列入学制。据《周礼》记载，

---

① 张政烺．试释周初青铜器铭文中的易卦．考古学报，1980，4

“（保氏）掌谏王恶，而养国子以道，乃教之六艺：一曰五礼，二曰六乐，三曰五射，四曰五御，五曰六书，六曰九数”。所谓“六书”，即殷商时期的“书学”，属识字教育的内容。东汉许慎《说文解字·叙》说：“周礼，八岁入小学，保氏教国子先以六书。”就是说，“六书”是“国子”入学后最先学习的内容。

此后，汉字教学就成为历代文化活动的重要内容，成为中华文明史的重要组成部分。秦承周制，是西周文化的直接继承者。当朝重臣丞相李斯、中车府令赵高、太史令胡毋敬亲自主持编写童蒙识字课本，可知秦王朝对识字教育的重视。汉代统治者不仅重视汉字教学，还在法律中作出具体规定。据《汉书·艺文志》记载：“汉兴，萧何草律，亦著其法，曰：太史试学童，能讽书九千字以上，乃得为史。”从此，识字教育成为全社会的自觉行为。汉时已经有了类似今天小学的“书馆”。王国维说：“汉时教初学之所名书馆，其师名曰书师，其书用《仓颉》《凡将》《急就》《元尚》诸篇，其旨在使学童识字习字”（《观堂集林》卷四《汉魏博士考》）。崔实《四民月令》说：“正月，民事未起，命成童以上入大学，学五经……砚冻释，命幼童入小学，学书篇章。”“小学”即书馆，“篇章”即童蒙识字课本。幼童即开始进行识字教育，历代无不如此。宋陆游诗曰：“儿童冬学闹比邻，据案愚儒却自珍。授罢村书闭门睡，终年不着面看人。”自注曰：农家十月乃遣子入学，谓之“冬学”；所读《杂字》《百家姓》之类，谓之“村书”（转引自钱大昕《十驾斋养新录》，见中华书局《四部备要》本第11页）。描绘的虽是南宋农村儿童习字学书的景象，但在整个封建社会中，也非常具有代表性。

二、科学的童蒙识字课本

由于我国的汉字教育起源早，因此童蒙识字课本的编纂也非常早。我国有史记载的最早的识字课本是西周宣王时的《史籀篇》。《汉书·艺文志》说："《史籀篇》者，周时史官教学童书也。"据专家研究，该书按字的意义关系编排而成，四字为句，二句为韵，以便学童习诵。这也是世界上第一部童蒙识字课本。秦始皇统一中国后，用小篆统一天下文字，李斯等人编《仓颉篇》、《爰历篇》、《博学篇》(后人称之为"秦三仓")，作为推行小篆的范本，也是童蒙识字课本。西汉初期，闾里书师将这三本书进行删改合并，去其重复字，合编为《仓颉篇》，用以教学童识字。西汉中期到东汉，识字课本的编纂达到高峰，著名的即有：司马相如的《凡将篇》、李长的《滂喜篇》、史游的《急就篇》、扬雄的《训纂篇》，等等。直到许慎《说文解字》出，识字课本的编纂势头才渐趋平息。魏晋以后，直到宋代，楷书成为社会通用字体，陆续出现了几部有影响的楷书识字课本。最有影响的有：南朝梁周嗣成奉敕编成的《千字文》，宋王应麟等编的《三字经》，宋佚名编的《百家姓》。三书后被合称为"三、百、千"，成为几百年使用不衰的识字课本，一直使用到清末，直到今天仍被人念念不忘。这些课本是汉字教育史上的宝贵财富，具有很强的科学性。主要表现在：

第一，选录常用字。如何在最短的时间内，获得最高的识字效率，使学习者尽可能快地获得阅读能力，汉字教学的内容是至关重要的。因此，识字教材选录的必须是常用字。古代童蒙识字课本在常用字的选择上非常成功，其重要标志是：每个时代的课本选字都能被后代的课本所传承。周宣王时的《史籀篇》，成为秦代识字教材的底本。据班固所云，李斯等编《仓颉篇》等书，"文字多取《史籀篇》，而篆体复颇异"；到东汉建武时，《史籀篇》仍

在，只是“亡六篇矣”；秦代《仓颉篇》仍被汉初闾里书师采用教学童识字；“三、百、千”流传在汉字教学中几百年而不衰。这些事实都说明，历代字书对常用字的选择是科学而成功的，其中的绝大多数字种都属社会基本用字，并不随着时代的变迁和时间的推移而被废弃。

第二，收字数量科学适用。汉字数量大，而且，作为一个开放系统，在几千年的历史中，汉字的数量总是在或快或慢地增加着。但是，汉字总数虽然增加，社会常用字种和字量却是基本稳定的。汉字初级教学字量应当与社会常用字量大致持平。古代的童蒙识字课本在字量的确定上为我们提供了范例。从《仓颉篇》到“三、百、千”，每部课本的收字量约在 1 000～2 000 左右，同时并用的成套识字课本的总字数约在 3 000～4 000 左右。据《说文·叙》，秦“三仓”经西汉闾里书师整理后，删去其重复之字，“合六十字一章，共五十五章”，共计 3 300 多字，则各篇约 1 000 多字。西汉初期，儿童识字仍以秦三仓为教材，识字量为 3 000 多字。西汉中期以后，汉代人开始自己编识字教材，《训纂篇》、《滂喜篇》各 2 400 字，史游《急就篇》2 100 多字，流传到清末的“三、百、千”，一共 2 000 多字。用当代汉字字频统计和社会常用字的研究成果进行验证，2 000～4 000 字几乎是中国历代社会常用字的恒量。周有光曾对当代各家对现代汉字字频的统计结果进行了分析归纳，抽象出了“汉字出现频率不平衡规律，也叫汉字效用递减率”①。这个规律可表述为：最高频的 1 000 字，覆盖率大约是 90%；以后每增加 1 000 字，大约提高覆盖率 1/10。根据汉字效用递减率，3 800 个字种覆盖率为 99．9%。我国历代字书的收字量，与当代人使用先进的科学技术统计出的结果若合符节，这充分体现了古代字频研究和社会常用字统计的水平。历代文学

① 周有光．中国语文纵横谈．北京：人民教育出版社，1992．156

名著不重复字的用字量也印证了历代识字课本体现出来的社会常用字的恒量。试比较：

书　名	成书年代	总用字量	不重复用字量
尚书①	先秦	24 538	1 941
易经	先秦	20 991	1 583
诗经	先秦	29 646	2 936
论语	先秦	15 918	1 382
春秋三传	先秦	245 838	3 912
淮南子②	西汉	133 827	4 208
战国策	西汉	122 529	2 774
说苑	西汉	112 254	3 271
孔子家语	东汉	57 423	2 566
全唐诗：张籍卷③	唐代	25 502	2 465
全唐诗：白居易卷	唐代	189 154	4 750
杜甫诗④	唐代	未详	3 562
三民主义⑤	近、现代		2 134
毛泽东选集 1～4 卷	现代	660 273	2 981
倪焕之	现代	138 000	3 039
骆驼祥子	现当代	107 000	2 413

说明：①先秦典籍字数统计摘自王凤阳《汉字学》；　②西汉、东汉著作字数统计据台湾商务印书馆《先秦两汉古籍逐字索引丛刊》；　③张籍、白居易诗用字统计据现代出版社《全唐诗索引》；　④杜诗和毛选字数据张朝炳《毛泽东选集用字的字数、次数按音节分布的情况》，载《中国语文》1980 年第三期；　⑤《三民主义》、《倪焕之》、《骆驼祥子》用字统计摘自范家莺等《现代汉字学》。

由上表可以看出，我国历代各类著作所用不重复字数很少有超过4 000字者。这说明我国历代的社会常用字有一个恒量，这就是在3 000字上下。因此，尽管史籍语焉不详，我国古代围绕汉字教育所进行的字频研究和常用字研究，已经达到了很高的水平。这也从另一个角度证明了我国童蒙识字教材编纂的科学性。

三、对识字方法的研究和探讨

我国古代的识字教材，在字种的排列上并非杂乱无章，而是有一定顺序的，后代字典所用的分别部居、据形系联、据义系联的排列方法，在这些字书中都已经萌芽。这体现了编纂者们文字学方面的素养，更为可贵的是，体现了古人对汉字教学法的探索和研究。

第二节　我国当前的汉字教学

我国是一个文明古国，自古有重视教育的传统。历代统治者对汉字教育尤为重视。我国大量的识字课本，编纂者多为国家重臣，或为主管文化的官员，如周代的太史籀，编写秦“三仓”的李斯、赵高、胡毋敬，汉代的扬雄、司马相如、李长、蔡邕，南北朝时的周兴嗣，宋代王应麟等，都是一时之名臣。新中国成立以后，党和政府把汉字教学提到重要的位置，投入大量的人力和物力进行研究与实践。旧中国，享受汉字教学的只能是极少数人，建国后，汉字教学的范围得到前所未有的拓展，原有的基础汉字教学得到空前发展的同时，又开创了汉字教学的新领域。围绕着汉字教学，教学常用字研究成为汉字研究的重要领域。

一、基础教育中的教学字表研究

基础教育中的汉字教学是汉字教育最重要的领域，它关系着全民文化素质的改善与提高，关系着民族的前途和文化的继承，因此，它也是汉字教学中最古老的领域。我国最早的识字课本就是为汉字基础教育而编写的。自宋代以来，“三、百、千”几乎成了全国通用的识字教材。进入20世纪以来，特别是“五四”以后，社会生活发生了重大变化，新式小学代替了旧私塾，迫切需要以新教材代替旧的识字课本。随着社会生活的变化，随着白话文代替文言文，社会常用字也由古汉语用字向现代汉语用字转变。这种转变，必然反映在小学教材中。因此，重新确定小学识字的字种和数量，便成为急需解决的问题。自20世纪到现当代，基础教育用字的研究从未停止过。其研究方法是，采用科学的统计方法进行汉字的使用频率的研究，将高频字作为教学用字。我们现代最早采用科学方法进行字频统计的是陈鹤琴，他为了研究语文教育问题，对语文课本及通俗书报的用字作了统计分析，从55万字的中文资料中选定4 261个字种，编成《语体文应用字汇》，于1928年由商务印书馆出版。我国现代最早的小学用字研究是1930年12月民智书局出版的王文新著的《小学分级字汇研究》。作者把收集到的小学生作文、小学教科书用字和我国第一部现代汉字字频统计著作《语体文应用字汇》所收字进行汇总统计，得出字种5 364个，根据识字要求，淘汰了1 565字，其余3 799字就作为小学识字用字。作者把这些字分为初小、高小共六个年级，并确定各年级的识字量和具体字表。其各年级的识字量的分布为：

初小：一年级541字，二年级558字，三年级654字，四年级793字；

高小：五年级622字，六年级631字。

近年来，为了加强识字教学的科学性，从而缩短识字教学的时间，提高识字教学的效率，对于识字量的确定、字种的选择，都越来越受到重视。除了上面提到的周有光总结出的“汉字效用递减率”之外，有关专家从语文教学的角度，对汉字的频率作了更精确的统计。据专家研究，从汉字的覆盖率来看，掌握 500 多个常用字，覆盖率就高达 80%；掌握 900 多个常用字，覆盖率能达到 90%；掌握 1 350 个常用字，覆盖率则达 95%；掌握 2 400 个常用字（也有人统计为 2 500 个），覆盖率就达到 99%。在汉字教学中，字种的选择应以科学统计为依据。一方面，识字的目的是为了阅读和写作，对于读和写来说，汉字的覆盖率越高越好，读写时接触到的字种中，已知的字种占的比率越高越好。另一方面，所选字种在读写时复现率越高，识字巩固率也越高。由此可见常用字种的选择在识字教学中的重要性。近年来，语文工作者和有关专家加强了这方面的研究。

1983 年，《教育研究》第五期发表了张卫国的《小学语文用字研究》，其中《小学语文用字表》收字种 3 071 个。

1985 年，北京语言学院语言教学研究所对 1978—1980 年出版的全国通用的中小学语文课本的全部词汇作了统计与分析，编成了“满足中小学语文教师、学生、学生家长和其他各类学习汉语者的迫切需要”的《常用字和常用词》。其中有《按出现次数多少排列的常用字表》，收字 1 000 个。作者统计，“中小学语文课本用作统计资料的全部篇幅，有近五分之四是用 1 000 个高频汉字写成的。”其中，前 500 字的覆盖率达到 69.98%。

1988 年，语文出版社出版国家语委汉字处编的《现代汉语常用字表》。该表由国家语委和国家教委联合发布，其目的是“为了适应语文教学和其他方面的需要”，收常用字 2 500 个，次常用字 1 000个。这为汉字教学、为语文教材的编写提供了依据。

国家语委发布的常用字表是国家对中小学语文教学字种字量

的规定，属文件性质。对中小学语文教师来说，对汉字教学具有工具书性质的、最方便实用的常用字表是娄警予、王贵文编著的《现代汉语常用字表》。《字表》收入1987年北京教育出版社出版的《语文基本知识小丛书》(叶苍岑主编）中。该书作者是两位语文教师，他们编写《字表》的目的很明确：为中小学语文教学服务，“改变中学汉字教学‘一笔糊涂帐’的状况”。该“字表有一个突出的特点，就是实践经验、科学分析、仪器运用相结合。他们先凭学汉字、用汉字、教语文的丰富的实践经验，从可靠的字典辞书中选出他们认为应当列为常用字的字表，然后从许多个侧面对初步的字表进行客观的核查、检验——这其实就是科学分析”(张志公《序言》)，因此，可信度很高。同国家标准局用电子计算机统计所得的《信息交换用汉字编码字符集基本集》的常用字表3 755字相对比，有3 706字相同，所未收的49个字多属较生僻的专业用字、地名用字，等等。该字表共收5 623字，其中，小学教材中已出现过的字2 986个，中学语文教材出现过的1 496个，还有作者经过调查研究认为使用频率较高的、应该扩选入字表的665字。该字表的编写体例使它对语文教师具有工具书的作用。这就是，凡在小学语文教材的生字表中已经出现过的字，在字的下面标出该字在哪一册中出现；凡在中学语文教材中出现过的字，在字的下面标出出现在哪一册、哪一课中首次出现。这就为中小学教师系统地进行汉字教学提供了方便。中学教师可以从字表上看到中学生在小学阶段共识了多少汉字，并据此检查他们对小学阶段所识汉字的掌握情况；还可以从字表中看到小学阶段教学的每个汉字，分别出现在哪一册教材中，进而查阅小学教材，把两个阶段的识字教学连成一个整体。例如：

从《字表》中可以看到，“哀”出现在小学教材第五册中，从五册生字表中可以进一步查到，“哀”出现在第二十八课《寒号鸟》中：“崖缝里冷得厉害，寒号鸟冻得直打哆嗦，悲哀地说：哆

罗罗，哆罗罗，寒风冻死我，明天就垒窝。”“就在这严寒的夜里，喜鹊在温暖的窝里熟睡，寒号鸟却发出最后的哀号……”很显然，两处的“哀”字都表示“悲伤”的意思。这就是说，小学阶段只重点学了“哀”字的一个义项。那么，当中学教师在中学语文教材第二册第二课《古代英雄的石像》中，遇到“哀求”“哀吊”这两个词时，就应进一步教给学生“哀”的另外两个义项：在“哀求”中，“哀”有“苦苦地”的意思；在“哀吊”中，“哀”是“悼念”的意思。接着，当中学教师在中学语文教材第四册第十二课《二六七号牢房》中遇到“哀怜”一词时，还应该进一步告诉学生，“哀”在这里是“怜悯”的意思。至此，“哀”字的四个义项都教给了学生。教师可通过引导回忆和阐明各义项之间的联系的方法，使学生对“哀”字的含义获得完整的、清晰的认识，并能掌握其用法。

《字表》对中学语文教材中所出现的汉字的各个出处都作了标示，使教师可以有步骤地由浅入深地讲授每一个字。例如，从《字表》及其所附《中学语文教材生字表》中可以看到，“禀”在教材中主要出现过三次。教师查到这三个出处后，再根据“禀”字共有三个义项，就可以有步骤、有计划地进行教学了。在第一册第二十二课《鲁提辖拳打镇关西》“容奴告禀”一句中首次出现，义为“旧时指向上级或长辈报告”，教师教给学生这个义项的同时，还可补充说明“禀”的另一个义项是名词，指“旧时禀报上级的文件”，如“禀帖”。这样就一次教给学生两个联系紧密的义项。当第四册第九课《第二次考试》中出现“禀赋”、第六册第十九课《葛梅》中出现“禀性”时，教师就可以教给学生，这两个词中的“禀”都是“承受”的意思。这样有步骤、有计划地进行汉字教学，既可以保证学生对每个字的各个主要义项有全面的、清晰的认识，又可以在教学中分散难点，并使每个义项的讲解都紧密结合课文的内容，收到事半功倍之效。特别重要的是，教师对每个汉字都

实行有步骤、有计划、有联系地教学的方法，会给学生以启发，使他们懂得汉字各个义项间有内在联系的客观规律，从而掌握学习汉字的科学方法，提高自学的能力。

综上所述，《字表》最适合中学语文教师进行汉字教学，也适合小学语文教师使用，同时也适合学生根据中小学语文教材自学汉字。

该《字表》虽然出版在中小学语文教材改编之前，但是并不过时。这是因为：第一，《字表》所收字与《现代汉语常用字表》大致相符，是中小学生必须掌握的常用字；第二，由于《字表》对每个汉字在中小学教材中的出处都作了标注，因此，即使中小学教材有变动，《字表》及其所附的《中学语文教材生字表》，只要随之作相应的变动和修改仍然适用。

《字表》标示汉字在小学语文教材中的出处，以中小学通用教材小学语文编写组编写的全日制十年制小学课本（试用本）为准。标示汉字在中学语文教材中的出处，以 1982 年度全国通用的初、高中语文教材为准。现在全国中小学已经使用新教材，教师可以自己对《字表》进行补充。

《字表》虽然科学实用，但由于只印了 25 000 册，因此影响有限。鉴于中小学教材已经改编，编写新的、针对中小学语文教材的、对中小学汉字教学具有指导作用的类似的常用字表已是当务之急。

二、汉字扫盲教学及其教学用字研究

汉字扫盲开始于 20 世纪初期。封建统治阶级历来奉行“民可使由之，不可使知之”的愚民政策，决不鼓励“劳力者”学习文化，掌握汉字，因此解放前文盲占人口的绝大多数。鸦片战争中，帝国主义用火炮轰开了中国的大门，使中国沦为半殖民地。忧国

忧民的知识分子为此认识到了唤醒民众的重要性，提倡教育救国，开始了汉字扫盲运动。一些知识分子身体力行，编排扫盲用字表，并亲自到乡下，进行扫盲工作。扫盲字表的编写，是扫盲最基础的工作，也是提高扫盲效率的保证。我国最早的当时影响最大的扫盲字表是洪深编著的《一千一百个基本汉字使用教学法》，由生活书店 1935 年 11 月出版。该书扉页题字“献给推行识字运动的诸位先生”，揭示了其写作目的。作者选定 1 100 个字作为基本汉字，另外再加 250 个特别字，主张用这些字来“表达一切的心情和事实”，这实际上是限制用字。他说：“将来，一切政府的文告，法院的批判，学者的理论，报纸的社评，也许可以全用 1 100 个基本字和 500 个特别字来写，使得那凡是经过识字训练的民众自己都能看懂。”这段话进一步说明此书是为扫盲工作提供的。

洪深所谓“基本汉字”，不是根据字频来选定的，而是根据字义，根据表达的需要，选取每组同（近）义字中最有代表性的一字。洪深主张，写作时遇到基本字以外的字，就改换一个说法。例如，基本字中有“土”而无“泥”，需要用“泥”的地方就用“湿土”；有“地”而无“田”，“种田”要改说为“种地”；没有“媳”，“儿媳”要改为“儿子的老婆”；等等。

洪深的主张在当时影响很大。他所选的 1 100 个基本字，都是汉语中最常用的字，覆盖率可达 90%以上，他本人也身体力行，用这一千多字写出了很精彩的文章。但是，洪氏的主张是不懂语言学的外行主张，其致命弱点是颠倒了语言与文字的关系，用文字来限制语言，人为地改变语言来迁就被缩减的汉字，这是削汉语之“足”来适应汉字这个“履”。因此这个主张是注定要失败的。但是，洪深代表了当时知识分子对扫盲工作的热忱和探索；他的失败，也为语文工作者提供了借鉴：扫盲工作和其他语文工作一样，不能违背科学规律。

在旧中国，虽然进步的知识分子为扫除文盲进行了不懈的努

力，但是，广大民众在政治经济上没有改变地位，扫盲不可能取得显著成果。收效最显著的是共产党领导的扫盲运动。早在战争年代，伴随着大生产运动，根据地就开展了扫盲运动，形成了参加扫盲的热潮。建国以后，党和政府立即着手扫盲工作。1949 年 11 月，中央人民政府教育部设“识字运动委员会”，以推动识字教育和扫盲工作。1952 年 6 月，教育部公布扫盲常用字表，收字 2 000个，供编识字课本用。1953 年 11 月，扫盲委员会制订《扫除文盲标准》，规定农民识字 1 500 个，工人识字 2 000 个。此后扫盲工作一直常抓不懈，有专门机构负责，取得了重要进展。1984 年教育部公布，建国 35 年以来，全国共扫除文盲 1.5 亿人，文盲占总人口的比例由 1949 年的 80%以上下降到 1982 年的 23.5%。

改革开放以后，由于一个时期全民经商热潮的出现，扫盲工作出现反复，文盲率有所回升。据 1988 年人口抽样调查估计，我国现有文盲两亿左右，每年仍有 200 万左右的新文盲产生。根据这种情况，国家加大了扫盲工作的力度。1988 年 2 月 15 日，国务院颁布了《扫除文盲工作条例》。这是为了加强我国的扫盲工作，尽快完成扫盲的历史任务而制订的行政法规。其中规定了个人扫盲工作的标准，内容是：“农民识 1 500 字，企、事业单位职工、城镇居民识 2 000 字；能够看懂浅显通俗的报刊、文章，能够记简单的帐目，能够书写应用文。”1993 年，国家语委所属语言文字应用研究所研制了《扫盲用字表》（载《语言文字应用》1993 年第三期）。该表分甲、乙二表，甲表收全国统一的扫盲必写字 1 800 个，乙表收扫盲参考用字 200 个。上述文件是国家扫盲工作的国家标准和规范用字。

扫盲工作受到了世界各国政府的重视。1985 年，联合国教科文组织大会第二十三届会议，认为在 2000 年扫除文盲是一项特别紧迫的任务。联合国大会于 1987 年 12 月通过决议，规定 1990 年为国际扫盲年。

三、对外汉字教学及其标准

对外汉字教学是一个古老而新兴的领域。早在3世纪，就已经有了对外汉字教学活动。据史书记载，公元3世纪，汉字传到了朝鲜；此后，又相继传到了日本、越南等国，成为上述国家记录自己民族语的符号或辅助符号。汉字的对外传播需要对外汉字教学。因此，从理论上说，自我国与他国开始进行文化交流之后，对外汉字教学就已经开始。只是由于有关史料的缺乏，我们难以了解其详细情况。

解放以后，特别是改革开放以来，对外汉语汉字教学的需要显得日益迫切，并得到了迅速发展。据1989年统计，全国开展对外汉语教学的高等院校达到一百余所。制定科学的用字标准的要求，也提到议事日程上来。1990—1991年，国家对外汉语教学领导小组和北京语言学院汉语水平考试中心联合研制了《汉语水平词汇与汉字等级大纲》。这份大纲是我国初等、中等汉语水平考试的主要依据，也是我国对外汉语教学总体设计、教材编写、课堂教学和成绩测试的重要依据。大纲中汉字的分级与词汇的分级保持一定的联系。汉字大纲收字2 905个，分为四级：甲级字800个；乙级字804个；丙级字601个；丁级字700个。这2 905字中，有2 485字是《现代汉语常用字表》中的一级常用字。

第三节 汉字科学对汉字教学的指导

汉字教学是基础教育重要的组成部分，是扫盲教育和对外汉语教育的重要内容。无论汉字教学的对象是哪一类人，学会汉字、掌握汉字都是一项很艰巨的任务。因此，能否掌握科学的教学方

法，便成为能否尽快地、高质量地完成教学任务的关键。

汉字是表意文字，即根据它所记录的词义构拟形体。因此，从理论上说，语言中有多少词，就会有多少字形。但是实际上，由于社会用字中对多义共一形的现象在一定程度上的容忍，特别是由于同音借用，都限制了字形的无限增多。但是，汉字的数量依然很大，仅社会最常用字就有 2 000～3 000 之多。掌握这些汉字，对学习者并非轻而易举。从结构上看，汉字是形音义的统一体，其中语音的体现是隐性的，没有直接的、准确的表音成分。汉字结构复杂，又不能见字识音，给汉字教学增加了难度。但是，只要按照汉字构形的规律做到汉字教学的科学化，教学的难度就会减少，教学的速度和质量都会得到根本的改善。

识字是求知的前提，识字活动是一种终身的学习活动。各种识字教学尤其是小学识字教学，要为这种终身的学习活动打好基石；所以，我们不能把识字教学的目标仅仅设定为认识几个字，而应当从终身自我教育的长远目标来考虑其效果。从这个长远目标出发，识字教学不但必须遵循汉字自身的规律，而且要使学习者在不言中感受到汉字的表义特点、汉字的构形规律性和汉字的系统性。而且，识字教学也只有遵循汉字构形的规律，才能在保证质量的前提下提高速度。

下面，我们举例说明汉字科学对汉字教学的指导。

一、根据汉字的属性来确定识字教学初期积累的字表

识字教学是分阶段进行的，每到一个阶段，教学方法和策略都要因积累的不同而发生变化。提高教学质量的关键是选择好初期积累的字，也就是选择第一批字来突破"零"。这一批字必须是比较容易的，是能够更有效地带动以后的学习的。确定这一批字

必须根据汉字的属性，而且要根据汉字的多种属性综合确定。这些条件是：

(1) 结构相对简单，即构件一般不超过2个；

(2) 构意明晰度高，即不含仅有代换功能的记号构件；

(3) 构字频度高，特别是作表义构件的构字频度高，有利于带动第二阶段的教学；

(4) 一般先教书写自由语素的常用字，选择书写不自由语素的单字时，应同时选择经常与之构成双音词的另一个字进入初期积累字表；

(5) 初期积累字对应的词，应是6～7岁儿童口语中已经会说的，也就是音与义已经被学习者把握了的；

(6) 适当选择虚词，以便组句。

要同时满足以上六个条件，必须运用汉字构形学所提供的汉字属性——包括构形属性、职能属性和字用属性，进行综合的择定。

二、依据汉字构形规律，科学地讲解字理

汉字属表意文字系统，科学地分析和讲解汉字即通过构形讲解它的构意，可以帮助我们理解和记忆字所记录的词义。但是，有些人存在一种错误认识，认为现代汉字已是一堆没有理据的符号，可以随便讲，只要能让学生记住就行。实际上，汉字学是一种科学，汉字的讲解必须是科学的。不科学地讲解汉字，会造成学生的思想混乱，而且也不利于他们对汉语汉字的进一步学习。

汉字构形的知识是用来科学地分析汉字的。在讲解汉字时，必须遵循以下几个原则：

(1) 汉字是由构件组成，其构意的体现是通过构件来实现的，讲解汉字，首先不能将其构件讲错。如："福"是从"示"、"畐"

声的形声字，有人把它拆成“衣”、“一”、“口”、“田”，把它讲成“一口人有田种，有衣穿，就是福。”且不说这种讲解在构件拆分上的错误，单是把“示”讲成“衣”，已经造成了一种大混乱。再如：有人把“悲”中的“非”讲成“韭”，说“悲”就是“心里像长了韭菜一样悲哀”。

（2）汉字是由构件组合而成的，构件进入构字后，便具有了构字功能：或示音、或表义、或表形、或起区别标示作用。讲解汉字必须以构件在构字过程中的实际功能为依据，讲错了或曲解了构件的功能，就会使整个字的讲解发生错误。例如：“韭”是一个独体象形字，小篆字形作“韭”，像韭菜丛生地上之形，它是以构件的表形功能来体现构意的。有人因为“韭”中的“非”与“非常”的“非”外形相似，就把“韭”拆分为“非”和“一”，把“韭”看作会意字，把它的意义和否定词“非”联系起来，讲成“不是（非）一根，而是一大片”。把表形构件“韭”讲成表义构件“非一”实在是生拉硬扯，牵强附会。

再如“饿”字中的“我”是示音构件，没有表意功能。但有些人把“饿”讲成“我要吃（饣），因为我饿了”，把“我”曲解为表意构件。这样就会造成一种误导，“俄”、“娥”、“峨”、“鹅”等从“我”得声的字就会被类推成“我的人”、“我的女儿”、“我的山”、“我的鸟”，将学生引人歧途。汉字是以形声字为主体的，示音声符和表义义符都对形声系统起归纳作用，把示音声符讲成表义义符，就会使整个形声系统的规律不能展现，造成人们对整个形声系统认识上的混乱。所以，我们不能为了一时方便而随便乱讲，造成讲了一个，乱了一片的严重后果，给识字者造成困惑。

（3）由基础构件组构成汉字，大部分是依层次逐级组构的，构意是逐级生成的，以集合的方式产生构意。在解释汉字时，必须按它们组合的实际状况来进行，既不能把层次结构讲成平面结构，也不能把平面结构讲成层次结构，否则就会发生错误。而人们常

见的错误就是不懂得汉字结构层次生成的道理，把汉字的所有构件看成同一个层次的东西。例如，有人把“温”字讲成“太阳照在器皿里，水就变温了”。这种讲法是错误的，因为“温”是一个层次结构的字，“日”不是“温”的直接构件，不直接对“温”起构意作用，它的功能只是与“皿”构成“温”的示音构件“昷”，“昷”和“氵”再组合才构成“温暖”的“温”。而且，这个“日”的形源也并不是“日月”之“日”。

（4）汉字的构形是发展的，今文字中失去理据不能分析的字形，要通过传承线索进行复形。例如，现代汉字的构意多数仍可以分析：

“张”、“弛”从“弓”，它们的本义是指把弓弦拉紧和把弓弦放松。

“引”从“弓”，本义是开弓的动作。

“弯”、“弧”从“弓”，它们表示的是弓背的状况。

“红”、“绿”、“紫”从丝旁，它们都表示某种染色。

“访”、“谏”、“谚”从“言”，因为它们的本义都与说话有关。

但是，现代汉字已经是经过长期演变的字形，不能直接分析的形体已比古文字多得多了。因此，在分析现代汉字时，需要作一定的工作，使形义统一的条件具备。对字理的分析来说，为由于字形演变而不能直接分析构意的字复形，是一项十分重要的工作。复形就是通过对字形的历时认同，沿着汉字发展的历史脉络向上溯源，恢复保留着构意的字形，通过早期的字形使形义关系显现出来。

第一章已经指出，汉字由起源而渐进为现代楷书，其间经历了很多的发展阶段。就形体的外观来说，不同历史时期的汉字存在较大差别，但各历史阶段又是相互联系、一脉相承的，汉字字体的演变有一定内在规律和一定发展趋势的。我们在复形时，只有依照汉字形体演变的规律和趋势去做，才能不主观而有说服力。

例如：

“壤”、“穰”、“瓤”、“镶”所从的“襄”字，构意难明，复形时先找到小篆，《说文》作襄，从衣，中间的形体仍不明白。再找到金文，发现它的过渡字形有[illegible]形，原来小篆字形中的两个“口”字是祭器的两耳的讹变。这个字的构意是以在土地上支起祭器攘除恶鬼来表示“攘除”义。

（5）分析字理要符合总体的构形系统。汉字不仅是历史的现实，而且是共时的系统。前面已经说过，汉字的构形是成系统的，它们相互之间发生着网络的关系，考察汉字的形与义，讲解汉字的形与义，都不能仅对单个汉字从主观的猜测出发，随意地作出结论，以免讲了一个，却与另外一个甚至一批发生矛盾，这是不科学的，也无法说服人，更不利于教学。

有一种被称作“流俗文字学”的主张，他们认为汉字是一种拼形拼义的文字，现代汉字的字理大都可以用会意的方法来解决，而且可以不顾字形演变和结构的特点用主观联想来任意讲解。前面所举的例子大都是从这种主张中挑出的。这种主张在理论上有以下三方面的错误：第一，它违背了汉字本体的事实，忽略了现代汉字大约有 15%左右的字形，即使在直接构件组合（即一级组合）层次上也已经无法分析字理，硬加分析，必然走向荒谬；第二，它否认或根本不知道汉字形义的统一，也就是构字的理据早已是客观存在的历史事实，只能解释，以求其相对准确，而不能编造杜撰；第三，它对汉字是一个具有不同程度严密性的体系视而不见，否认或不知道汉字的构形系统是建立在形声系统的基础上的，每一个字在这个系统中都有自己适当的位置，具有纵横两方面的字际关系，对任何字的讲解都会涉及其他的字，是不允许有随意性的。

汉字教学长远的目标，决定了必须在教学中体现对汉字正确的理解；汉字自身的系统性和结构的规律性，给汉字教学提供了

增加形象性和趣味性的足够条件。教师的创造性必须也绝对可以在科学的基础上体现出来，这就需要提高汉字科学的水平，提高对各种不同说法的识别能力。

三、根据不同教学阶段和汉字不同的属性，选择不同的教学策略

我们主张“教无定法，殊途同归”，识字教学的具体方法和切入点可以选择。但是，汉字的科学规律是不能违背的，这就是“教有定则，勿离勿违”的原则。汉字教学的方法不应当是一元化的，不但在每个阶段是不同的，对属性不同的字，也要选择不同的策略来教。例如，初期积累阶段，在教学传承独体字时，可以采用古文字作背景，沟通物象、古代象形字和楷书，而对粘合独体字就不宜用这种办法；又如，对理据明晰度高的会义合成字，可以以个体为单位，多用字理讲解法，而对形声字，则宜于用系统归纳法；再如，对表音度比较高的形声系统，可以利用声符系统编成韵语成批讲解，而对表音度很差的声符系统，就不宜于采用这种办法。

汉字科学是一门内容非常丰富的科学，教学方法的多样化，必须建立在对汉字科学深广的把握上，汉字科学的基础理论尤其需要认真普及。汉字构形学要解决的是汉字科学中最基本的问题，希望它对汉字教学起到应有的作用。

第七章
汉字与信息处理

我们正处在信息革命的新时代，各种信息在人类社会生活中的各个领域正起着越来越重要的作用，信息的存储、加工和传递的手段日益现代化，计算机技术得到了更为广泛的普及和使用。语言文字是信息的最主要的载体，因此，语言文字和信息革命有着不解之缘。信息革命中的绝大多数信息都要通过语言文字这一载体来储存、来加工、来传递。这样，仅有四十多年历史的计算机技术便向已有六千年左右历史的汉字提出了挑战。

世界上第一台电子计算机 1946 年诞生于美国，之后的短短四十多年里，计算机技术已经渗透到国民经济的各个领域，作为 20 世纪最伟大的发明，计算机已经获得了非常广泛的应用，特别是 INTERNET 的出现更使其应用达到了前所未有的高潮。从某种意义上说，计算机的发展速度和应用水平，已经成为人类进入信息化社会、国家走向现代化的一个重要标志。

但是众所周知，计算机在我国的推广应用要比西方发达国家落后许多年，除了经济基础和科技发展条件的制约之外，一个很重要的原因就是汉字进入计算机要比拉丁文字困难得多。因此，如何使计算机具备汉字处理能力，便是一个亟待解决的技术难题。

我国在 60 年代末期就开始对汉字信息处理进行探索和实践，在经过三十多年不懈努力之后，已经在许多研究领域取得了突破和进展。从 60 年代的汉字电报译码机、70 年代的新型的汉字输入输出设备，到八九十年代的各种基于个人计算机的汉字信息处理

系统以及用微处理机控制的汉字终端，汉字信息处理的研究成果，已开始在事务管理、企业管理、科技情报检索、机器翻译、人机对话、汉字编辑排版、咨询服务、公用事业、计算机辅助教学、计算机通信网络、办公自动化等方面得到推广和应用，这些都必将对我国的现代化建设产生深远的影响。

如果说在计算机汉字信息处理初期，汉字如何进入计算机是亟待解决的技术难题的话，那么在进入 90 年代以后，随着汉字信息处理技术应用的不断深入，汉字如何规范地进入计算机就成为主中之主、重中之重的难题了。因为计算机的日益普及，已经使得计算机的应用成为一种文化——计算机文化，这种文化在切合社会发展的同时，伴随着汉字在计算机上多种不规范的使用，已经产生了一些负面影响。为了消除这种不良影响，就应该在汉字规范地进入计算机或在汉字信息处理的规范化方面采取有效措施，加以解决。否则，技术虽然进步，文化却遭到破坏，这是得不偿失的。

第一节　汉字信息处理的过程

为了了解汉字与信息处理的关系，首先有必要了解已经成熟的汉字信息处理的过程。通常汉字信息处理的过程分为三个阶段：①汉字信息的输入；②汉字信息的处理；③汉字信息的输出。三个阶段的具体处理过程可由图 7-1 描述。其中汉字信息的输入是通过各种输入设备完成的，汉字输入设备及其设备驱动程序负责把汉字的外部码转换为处理系统识别的内部码。目前除了语音输入、字形输入和键盘编码输入外，还有通过通信设备和文件交换设备把汉字信息从一处传输到另一处。汉字的信息在机器中（指处理系统）通过特定的程序（软件系统）进行加工，如编辑、排

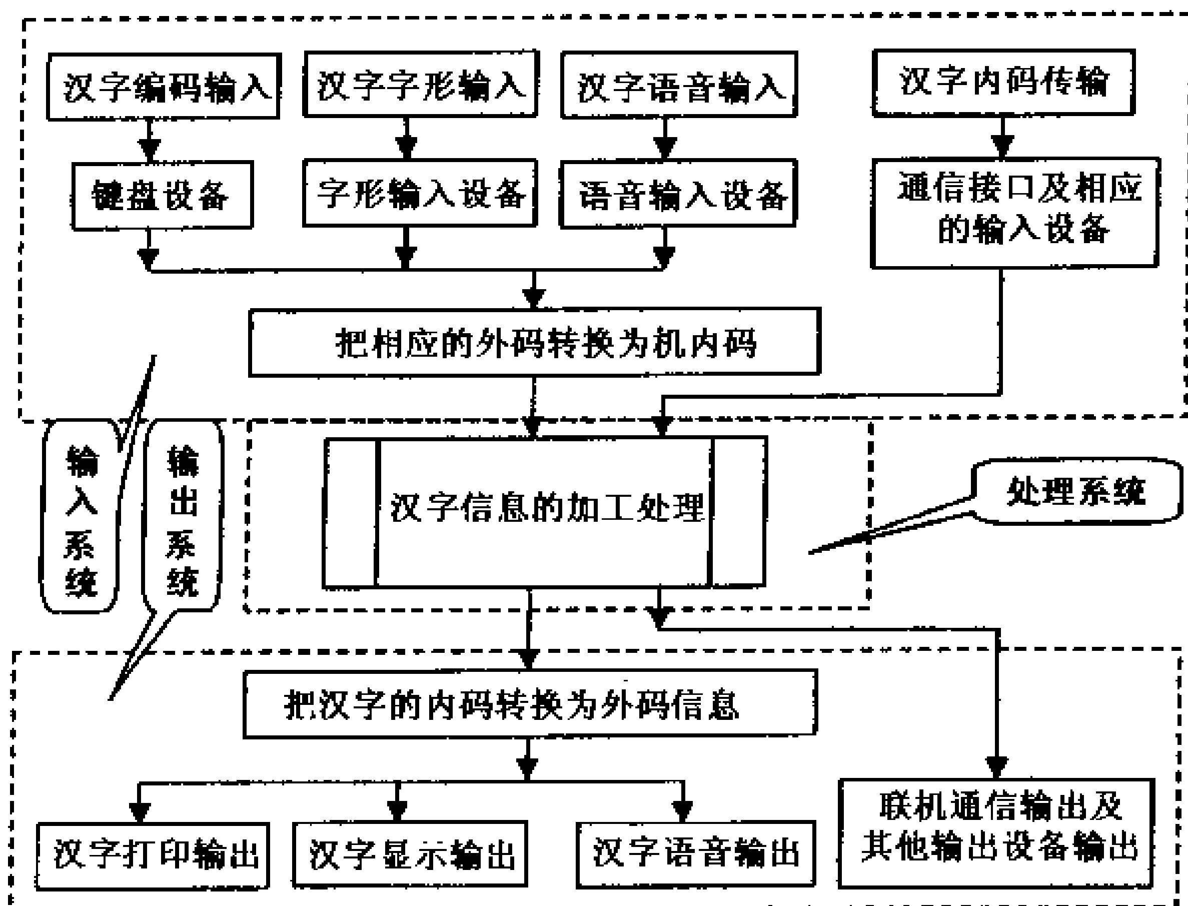


图 7-1 汉字信息处理的三个阶段

版、排序等，最后按照用户的要求进行输出。汉字的输出是把汉字内码转换成汉字外部字形和字音信息的过程。输出设备驱动程序通过汉字语音库、汉字字形信息库（简称“字库”）等加工完成后的汉字信息输出到显示输出设备、打印输出设备、语音输出设备上，同时，也可通过通信设备（如 FAX、MODEM 等）、文件交换设备把处理、加工完后的汉字信息输出到其他地方。

上述三个阶段中，汉字信息的处理与汉字信息的输出对于一般用户而言，无需了解内部实现机理，只要能够应用即可。而与用户最直接相关的就是如何将汉字输入到计算机中，目前解决的方法有三种：第一种是用键盘将汉字输入计算机，称为汉字的键盘输入；第二种是用文字字形识别的方法将汉字输入计算机，称

为汉字的字形识别输入；第三种是用语音识别的方法将汉字输入计算机，称为汉字的语音识别输入。

汉字字形识别输入就是通过与计算机联机的文图扫描装置，抽取汉字字形特征使印刷在纸上或写在纸（或介质）上的汉字字符，由计算机进行辨识后形成汉字内部码的过程称汉字识别(CCR Chinese Character Recognition)。目前使用得最多的扫描装置是光学字符阅读机(OCR Optical Character Reader)，因此汉字识别输入也称为OCR识别输入。汉字识别在学科上属于模式识别和人工智能范畴，在应用上是汉字信息处理系统的一种高速自动输入方式，它也是新一代计算机视觉智能接口的一个重要组成部分。汉字识别又分印刷汉字识别和手写汉字识别，后者又分为脱机识别和联机实时识别。

汉字的语音输入是通过与计算机连接的传声装置，利用语音分析技术，由计算机进行辨识后形成汉字内部码的过程称汉字的语音识别。汉字语音识别在学科上属于模式识别和人工智能的范畴，在应用上属于汉字信息处理系统的一种人机交互自动输入方式，它也是新一代计算机听觉智能接口的一个重要组成部分。

在汉字进入计算机的三种方式中，以汉字的键盘输入使用最广泛，相应的软件处理技术也最为成熟。汉字的键盘输入又可分为两类：一类是用字母、数字或符号将汉字编成代码的间接输入；另一类是以整字从大键盘上直接输入汉字。目前多用前者，后者极少使用。因此，汉字的键盘输入通常是指汉字的编码间接输入，汉字输入码就是为输入汉字而按照一定的规则，对指定的汉字集中的元素编制相应的代码，简称汉字编码。而汉字集元素映射到其他字符集元素的一组完整规则，则称为汉字编码方案。伴随着计算机中文操作系统的问世，汉字键盘输入技术在一段时期里发展非常迅速，如雨后春笋般地出现了各种汉字编码方案，形成了万“码”奔腾的局面。

数以百计的汉字编码方案和实用系统的出现，一方面体现出这一领域百花齐放的学术气氛，使汉字进入计算机之后成为比较优秀的信息载体，从而对中文信息处理现代化和普及计算机教育起到了良好的推动作用；但另一方面，由于各种编码方案（这里主要指根据汉字的字形特征进行编码输入的“形码”）在对汉字的拆分、部件的选取和归类上，见仁见智、自成体系，有的编码方案甚至从自己的需要出发，对汉字任意“肢解”，从而给计算机应用、语言文字规范、计算机教育和识字教育造成很大困难。因为没有明确的标准和规范，这种被动局面持续达十余年之久。

第二节　汉字编码及其评价

没有明确的标准和规范可循，原因是多种多样的，一方面权威机构对信息时代的迅速到来缺乏足够的理论思想上的准备，另一方面在此研究领域认识上的不统一，缺乏对这一领域比较全面客观的认识。

受 1986 年权威机构组织实施的第一次汉字编码方案评测指标的影响，许多编码方案盲目地宣传自己编码方案的输入速度，在一段时间内，编码作者们纷纷组织录入竞赛并报道键入速度，用户购买时也倾向认为键入速度越快的方法越好。走向极端之后，有人甚至不切实际地公布过 500 字以上/每分钟的键入速度。实际上，不同的人、不同的工作性质有不同的要求，作为一般用户对输入速度只有一个最低的要求：30 字以上/每分钟，而第一位的要求则是希望编码方案越易学越好，越不易忘越好。500 字以上/每分钟的键入速度即使有，对大多数用户来说也是用不上的。事实上，有些输入方案提高输入速度是靠减少重码和增加非规则编码，即牺牲易学性来实现的。这种现象与权威机构的导向有关，纵观

国家技术监督局制订的《通用键盘汉字编码输入方法评测规则》的19项静态与动态指标的评测项目中，直接和速度有关的就占了15项。导致这种结果的原因有二：第一，在汉字信息处理初期，使汉字进入计算机是主要目标，而汉字编码的规范性及易学性当时并未作为主要因素来考虑，同时能够量化且易于量化测定的指标只有速度；第二，计算机并未全面普及，汉字输入一直是少数专职操作员的专业技能，面衡量这种技能的高低只是单纯考虑输入速度并不考虑其他因素，例如使用者所耗费的训练时间、不规范行为模式的形成、语言文字规范的丧失等均被忽略。

有专家对此种现象迅速作出过反映，北京语言文化大学张普教授的《走出汉字键盘输入的三个误区》一文对汉字键盘输入的三种倾向"重码率越低越好、速度越快越好、词库越大越好"提出明确的纠正，应该说在很大程度上促进这一领域的良性发展，但是如何更为全面地认识和评价汉字编码方案则是许多专家学者所思考的问题。

北京师范大学何克抗教授及其所领导的课题组在承担国家教委"八五"重点攻关课题《全国中小学教学用汉字编码规范与计算机输入系统研究》时曾提出过汉字编码的性能指标体系。

要实现一个理想的汉字编码方案及计算机汉字输入系统，应该满足下列指标体系：

1. 规范性

（1）规范的音码编码方案应符合国家颁布的《汉语拼音方案》；

（2）规范的形码编码方案应符合汉字结构规律和国家语委颁布的有关语言文字规范；

（3）如果研究的音码和形码方案希望在中小学也能推广使用，则还应符合语文教学规律和符合国家教委制定的《全日制小学语文教学大纲》中有关拼音和识字教学的要求；

（4）音码和形码方案的实现均应遵循国家技术监督局（前国家标准总局）发布的《信息交换用汉字编码字符集》的有关规定。

2. 易学性

（1）编码规则简单、明确，部件与键位的分布科学、合理，易学易记；

（2）对汉字的拆分规则和编码规则符合人们对汉字的认知过程，掌握编码方法所要求的知识能与中小学原有认知结构相吻合；

（3）学后不易忘记，间断使用时易于恢复；

（4）规则的二义性小，码表与编码规则的符合程度高；

（5）字量升级时，编码方法与规则不变；

3. 快速性

单纯追求高速度并不是一个理想编码方案的要求；但是既然作为一种汉字输入方法，如果达不到较高的输人速度，缺乏应有的效率，也是没有意义的。因此，快速性是汉字编码方案必须具备的一项基本指标。

快速性的客观指标分为两类：

（1）编码方案的指标评价：①静态字、词平均码长；②静态字、词重码率；③静态键位分布系数等；④不等长编码技术。

（2）软件系统的指标评价：①高频先见技术；②字、词混合输入（无需键位切换）技术；③重码自动区分、词组自动编码等智能技术。

4. 兼容性

在计算机系统中实现汉字编码方案时应考虑三个方面的兼容性即：

（1）与国际通用键盘兼容；

（2）与多种显示终端兼容；

（3）与通用操作系统兼容；

当今计算机软、硬件产品已日益趋向国际化，过分强调汉字

输入系统的专用性，不考虑与国际上流行的软、硬件相兼容的做法，不仅在国外无法应用，不利于国际间的交流，而且在国内也难以普及推广。

5. 一致性

编码规则应前后一致，贯彻始终，不能自相矛盾，同类结构的汉字应使用相同的编码规则，尽量避免随意性。

6. 完备性

信息处理用汉字集通常分三个层次：基本集（含国标一、二级汉字共 6 763 个），ISO-IEC 大字符集（包括基本集和辅助集，共 20 902 个汉字）和全汉字集（包括全部汉字约60 000个左右）。

由此可见，要实现一个理想的汉字编码，通常要满足快速性、易学性、兼容性、规范性、一致性及完备性等六个方面的要求。这些要求的重要性并不是均等并列的，而是针对不同的社会需求和使用对象而有不同的侧重。对于专职操作员和打字员，往往把快速性放在首位，在此前提下考虑易学性，而对其他要求则可放在次要地位；对于教师、作家、管理人员这类非专职操作员来说，显然易学性指标是首要的，在此前提下，考虑快速性，而对其他指标则可放在次要地位；对于某些专门应用领域，如大型图书资料检索、历史文献与古籍整理，以及古汉语研究等等，由于这类应用需要大字符集汉字的支持，所以必然要把“完备性”放在首位（如前所述，完备性是指编码方案不仅能支持基本集汉字的编码，还能以统一的规则支持大字符集乃至全汉字集的编码），在此前提下考虑快速性与易学性，而对其他指标则可不作要求；对于而向广大中小学生的教学应用来说，为了保持祖国语言文字的纯洁和统一，促进中小学语文教学质量的提高，在编码设计中显然要把规范性放在首要位置来考虑，在此前提下努力满足易学性、快速性、兼容性等指标的要求。总之，对编码方案的设计离不开对社会需求的分析，离不开具体应用领域和使用对象。但是，不管是

对哪一种应用领域和使用对象，易学性和快速性这两项要求总是最基本的和不可缺少的，而规范性则更是非常必要的。随着计算机应用的日益普及和深入，对编码方案的规范性要求更是迫在眉睫，如何做到规范，对哪一环节加以规范，则需要认真分析和研究。

第三节　汉字信息处理用汉字部件规范

尽管汉字编码方案有数百种之多，但是各种编码方案的设计过程却基本一致。不考虑音码，任何形码方案的设计实际上均包含四个环节，这四个环节环环相扣、彼此衔接构成一个完整的体系。这四个环节分别解决下面四个问题：①部件如何选取；②部件如何分类；③汉字如何拆分；④汉字如何编码。

部件的选取是指，如何从组成全部汉字的所有部件中挑选出适当的数量来作为形码的基本部件。

部件的分类是指，如何将第一环节中已选定的数百个基本部件进行合理分类，使之能较均衡地分配到数十个码元之上。之所以要对部件进行分类，是因为系统的兼容性要求采用国际通用键盘来实现汉字输入，通用键盘只有 26 个字母键可充当码元，加上数字键也只有三十来个，而基本部件却有数百个。因此，能否建立一个科学的部件分类系统对于一个形码方案是否易学易记，是否有较低的重码率，即对于解决易学性与快速性的矛盾有至关重要的意义。

汉字的拆分是指，如何按照汉字的结构规律与构字方式将一个合体字拆分为若干个基本部件和单笔部件。

汉字的编码是指，如何将上一环节中已拆分出的基本部件和单笔部件按照一定的规则进行编码，以便通过键盘键入相应部件

的代码，从而达到输入汉字的目的。

从上述形码设计的环节分析不难看出，部件选取和拆分的不同，正是各种形码方案区别的关键，而且从目前国内流行的形码方案中，对部件数量的选取，少的取50个左右（如五十字元和安徽声数码），中等的取100～200个（如全息码取100个，王码取150多个，郑码取170个，部首码取189个），多的取300个以上（如表形码和认知码）。对部件的选取原则也各不相同，如郑码的部件90%选自《汉字统一部首表》，其余10%是惯用的组字构件；部首码的全部部件选自新华字典的189部首；有的编码作者则主张形码部件应由通用部首中的独体字、呈相交笔画的偏旁、简化偏旁以及通用字表中三笔画以内的字等四部分组成；还有的编码作者干脆根据自己处理重码的需要任意选取部件。这也正是造成"万码奔腾"混乱局面的主要原因之一。那么现代汉字的构形到底有没有理据，汉字的部件到底能不能规范？答案是肯定的。

一、汉字科学对现代汉字部件拆分的指导

要想进行现代汉字的部件规范，首先必须树立汉字构形规律性的思想。

在汉字构形是否存在规律的问题上，存在着两种不同的看法。一种看法认为，古汉字是表意文字系统，每个汉字的构形都是以来自词义的字意为依据，是可以解释的。这种可解释性，称作理据或字理。现代汉字虽经过隶变、楷化阶段，但并没有改变它的基本性质，大多数字形仍具有理据，少部分字形构字理据虽不太明显，但形与理没有矛盾，完全可以参考历史分析出来。其中极少部分汉字字形与意、源发生矛盾的，又可以放到构形系统中进行优化处理。共时平面上的汉字具有内部的系统性。汉字与汉字之间存在着相互的联系，每个汉字在系统中具有自己的位置，并

受到相邻汉字的制约。所以，对于汉字构形的分析，从个体看，有字理作为依据；从总体看，有系统中的前后左右关系加以制约，是有客观规律的。部件规范应当是这种客观规律的忠实体现，而不是将一些人的主观意志强加给另一些人的随意行为。因此，在对现代汉字的部件进行规范时，必须定出符合规律、符合实际、符合优化的操作性条例。

现代汉字是指书写现代汉语的楷书字，它是由历史上的隶书、楷书直接演变而来的，从个体字形看，它又是自甲骨文以来各代字形直接或间接积淀的结果。从科学测查和量化的情况看，这种汉字保留理据的比例仍然很高，而且，这种理据可以分析或追寻。

例如：小篆中的“勹”（bāo）像人的两手曲形有所包裹的样子。凡从“勹”的字，多有圆曲、周遍、包裹、内聚等意思。而在现代汉字里，这一形体演变后，不像曲形包裹状了。但是，依靠组合和聚合，以群体作背景，仍能显示其构意。如包（婴儿在襁褓中，义为包裹）、匊（两手捧着细碎的米，义为掬起）、旬（日子经十而一度循环，义为十日）等。在组合的另一部分配合下，理据能够显现，甚至表现得更为明显。

又如，楷书的“四”在构字时是“网”的变形，它虽已失去网形，但却具有网意，为罟、罹、詈、置等提供理据；“矢”已不像一支箭，但却具有箭的意义，因而可以给矩、短、矮、矫等提供理据；“隹”已失去短尾鸟的形状，但却从语言中承袭了 zhui 音，因而可以给谁、椎、碓等字提供声音信息。

经过隶变时部件的粘合，加上受行书连笔的影响，原来的古文字基础构形元素产生了形体的粘连，有合二为一甚至合更多部件为一的现象。例如“辶”、“共”、“西”、“更”、“退”等字即如此。因此，在古文字的多部件合体字里，理据可以一直贯穿到最后一个层次。而在现在汉字里，理据大多数保留在一级部件的组合中，越到后面的层次，保留理据的数量越少。后来的偏旁、部

首分析法就是适应现代汉字这一特点而产生的。当然，这对分析和讲解现代汉字提供了部分依据。

由于书写的缘故，笔画趋于平直后，一些原来形体与意义完全不同的独体字，一旦进入构字，便发生了形体异化，变为同形。例如，青，小篆原从“丹”，楷书从“月”；朔、期，小篆原从“月”，楷书仍从“月”；服、俞，小篆原从“舟”，楷书从“月”。在分析理据时，必须承认“月”这一部件分别来源于“月”、“肉”、“丹”、“舟”等不同的独体字。

反之，原来形体与意义都相同的独体字，一旦进入构字，由于部位的不同和受相邻部件的牵连，又可能异化为不同形体。例如：在“尉”中，“火”异化为“小”；在“光”中，“火”异化为“业”；在“然”中，“火”异化为“灬”；在“赤”中，“火”异化为“𣥂”；在“黑”中，“炎”上部的“火”异化为“土”。但是，只要在分析理据时，承认“小”、“业”、“灬”、“𣥂”、“土”在“尉”、“光”、“然”、“赤”、“黑”中都同源，都是“火”的变体，字理便仍在分析中起作用。

汉字构形理据的客观性及其可追溯性，免除了部件拆分的主观随意性，决定了拆分正误的可辨性，也保存了它历史文化的本来面貌，所以，部件规范必须使依理拆分充分体现。

例如，两部件的“们”、“引”，三部件的“鸿”、“靴”（其中“江”、“化”为二次拆分），四部件的“啊”、“姿”（其中“阿”、“资”为二次拆分，“可”、“欠”为三次拆分），五部件的“器”（为一次拆分），六部件的“僵”（“畺”二次拆分）。这些拆分都是按理据、依组合层次的反向进行的。

又如，“赤”的字形原从“大”从“火”，“大”异化为“土”，“火”异化为“𣥂”，但“土”与“𣥂”仍明显区别为相接的两个部件，因此，参考字源将其拆分为“土”、“𣥂”，既合字形，又合字理。

但是，在下面两种情况下，依理据拆分是无法贯彻的：

(1) 由于现代汉字的简化与笔画的形成确实有一部分形体既不保留理据，又由于字形与意、源发生矛盾，而难以重现理据。

例如“甫”，原从“用”，“父”声，“父”即“斧”的古字，斧标志权力，所以，“甫”是男子的美称。在小篆中“父”与“用”已相交合为一形，无法分析。

又如“至”，甲骨文以一支箭（矢）射中目标会“到达”之意，《说文》小篆把箭头变成较大的弧形而失去“矢”形，解作“鸟飞从高下至地也”，对理据的讲解已很勉强。楷书将箭尾的上端拉平成横；曲处写作“厶”，下变“土”，外形成为三个相接部分（一、厶、土），与原来的理据毫无对应，难以重现理据。

再如，“舂”原像两手（廾）捧杵（午）在臼中舂米，是一次性合成的四部件字。隶变、楷化后，“午”与“廾”粘合成“𡗗”，本身是非字部件，不具音义，又不能拆分，无法分析理据。再加上“𡗗”与“春”、“泰”的上部同形而不同源，归纳字意也不可能，成为一些文字学家所称的“记号”，使它们所构的字不再能重现理据。

(2) 前面说过，一部分汉字到楷书阶段时，理据仅仅保留在第一构形层次里，以后的层次理据有丧失现象。但是，拆分下来的部件，其信息量和区别度都要符合信息处理的需要，部件的总体数量也要尽量控制，因此，部件在一次拆分以后，还需要再度拆分。

因为有这两种情况的存在，在有理据拆分之外，还必须补充以无理据拆分。无理据拆分仍然不能是随意的，必须依照汉字形体的特点和部件的系统性来进行。具体做法是：

理据应合而楷书分离的，如“冓”，原像交构连接之形，上下本不可分，楷书将上部与下部分别楷定，成为两个相接可分的部分；又如“朋”，甲骨文像两串相连的贝串，小篆像鹏鸟的羽翅，

都相连不能分，楷书以两个相离的“月”字构形。在这种情况下，字形有拆分的可能性，这就是拆分规则所说的“相离可拆”。但“可拆”不等于“必拆”，是否拆分，要看拆分后的部件有没有归纳的可能性，是否有利于部件系统的优化。

理据应分而楷书交织、粘合的，这里又分两种情况：

(1) 原有部件完全交融，形体完全成为笔画的直接组合。如“甫”、“史”、“吏”等，既看不出拆分的可能性，又没有拆分的客观依据，应视为新的独体字。

(2) 少部分现代汉字，本为古文字描写性的隶定字楷化而成。它们的构形与意、源本是一致的。例如“秉”，原取“手握禾苗”之意，“禾”与“彐”（“又”的变体）是相交组合；又如“東”，小篆取“日出东方”之意，“日”插在“木”中组合，“束”、“口”像捆绑状，与“木”相交等。这种穿插结构原是多体象形的遗存，楷书象物性消失后，部件交合的原因已无法解释，其实与第一种情况已没有什么两样。这时也应尊重字形，不再拆分。

以上两点就是拆分规则所说的“交重不拆”。制订这条规则，还为了在统一的原则下控制拆分的下限。一旦允许交重可拆，无论从什么地方拆，拆成几部分，都无法避免如本文开头所举之例那样的随意性。另外，“交重可拆”还会破坏基础部件集的系统性。

在有理据拆分中，字形和字理是一致的，因此属于依形拆分。在无理据拆分中，字形与字理发生矛盾而采取尊重字形的原则，因此也属依形拆分。尽量尊重理据而不违背字形，其目的是尊重历史而不复古，立足现代而合乎规律，这样做，既维护了汉字的历史传承性，又维护了汉字共时的系统性，使汉字信息处理用的部件在符合规律的基础上取得一致。

形与理的一致性，反映了汉字历史与现状的契合，而形与理的矛盾则反映历史与现状在发展中的变革。把形与理放到两个层次中去处理，目的是为了从现代的实际出发，同时又尊重历史与

传统。只有这样做，诸多矛盾在不同层次、不同维度之中才能化解，使部件系统达到优化。

自觉遵循汉字字身的构形规律性，同时因势利导地调整汉字的社会约定性，立足现代，尊重历史，许多问题才能得到合理的解决。这是汉字部件规范所显示出的更为深刻的意义。

二、部件规范的实施步骤与标准

理解了汉字构形的规律性以及明确了汉字部件规范的意义，那么接下来要做的便是规范部件的具体实施步骤。1997 年 12 月国家语言文字工作委员会颁布了《信息处理用 GB13000.1 字符集汉字规范部件》。本节从以下几方面就其相关内容来具体介绍：制订规范的可行性、制订规范的基本原则、汉字构形原理及相关的概念术语、部件拆分的原则和部件归纳的原则等。需要说明的是，本节及下节内容均选自国家语委《汉字部件规范》研制组“关于制订《汉字部件规范》的若干问题”论文。

（一）制订规范的可行性

(1)规范性是评价汉字编码及其他信息处理系统的重要标准，也是使用者，特别是中小学教育界的使用者，选择输入系统的首要条件。过去由于缺乏部件规范，使编码的制订者和计算机的使用者缺乏衡量与选择的标准，造成了很多混乱。因此，制订规范已经成为大家共同的要求，具有广泛的社会基础。

(2) 在有关 GB13000.1 字符集的字形规范、笔顺规范、基本术语规范等相关的规范中，有的已经制订，有的正与部件规范同步制订，这些都为部件规范的制订工作打下了良好的基础。

(3) 进入 90 年代以来，有关计算机编码技术和汉字构形规律的理论研究已取得了实用性成果，使部件规范的制订具有科学的

理论依据。

(4) GB13000.1 大字符集中的西文操作系统和汉字部件自动拆分软件及其他工具的研制成功，使研究、制订部件规范有了高效可靠的技术手段。

(5) 信息处理界与语言文字学界的长期合作，促进了彼此的交流、沟通与理解，使制订部件规范所需的学科交叉和理论交融有了比较成熟的条件，也为部件规范的制订准备了具有合作精神的专家队伍和专业人员。

(二) 制订规范的基本原则

中文信息处理用汉字部件规范必须在两个方面具有严密的科学性：首先，必须符合汉字构形规律；其次，必须在部件的信息量和系统性上符合信息处理的需要。只有在这两方面是科学的，才能避免在部件规范形成中易出现的主观随意性，使最终成果既具有规范意义，又具有应用价值。

部件规范所确立的是用以衡量中文信息处理工作规范性的基础部件及其使用规则。它在拆分下限和拆分与归纳原则上应有严格的限定性；而用于编码设计时应在不违背规范的基础上，给出自由创造的余地。这两方面应在《基础部件表》和使用规则中体现。

为保证部件拆分与归纳的科学性，必须在理论上和实际操作中处理好汉字形体与理据的关系、历史与现实的关系、普遍规律与特殊情况的关系。研制组对上述三种关系的基本认识是：我们所要处理的汉字是现代汉字，因此，必须从现代汉字的实际出发。但是，现代汉字是从历史汉字传承而来的，当现代汉字的字形规律不够清晰时，必须参考历史，从字源上更深入地了解其构形规律。汉字属表意文字，其中相当一部分现代汉字构形是可以分析理据的，另有一部分在参考字源后，仍可追溯其理据。在这两部

分汉字中，理与形是一致的，从理也就是从形。只有少部分现代汉字的理据不能直接分析，追溯历史后，理与形仍产生矛盾。当处理这些汉字时，应从现代汉字的实际出发，从形不从理。从总体看，汉字的构形是有规律的，又是成系统的。但是，由于现代汉字是变动不定的开放系统，GB13000.1字符集中的汉字又有一部分是日、韩专用汉字，必然会有游离于系统之外的特殊情况。对于这些特殊情况进行处理时，应尽量选择既不违背实际情况又适应总体系统的最佳方案。

正是基于以上认识而形成了制订部件规范的基本原则，即“从形出发，尊重理据，立足现代，参考历史”。

（三）汉字构形原理及相关的概念术语

1. 汉字结构分析

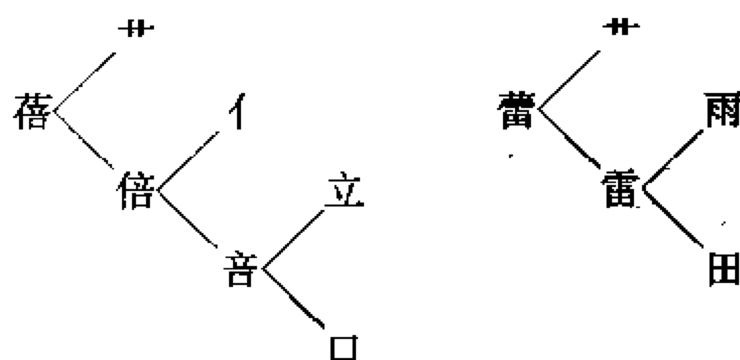
（1）结构理据。根据字源或参考字源，从汉字的部件组合中分析出的造字意图称结构理据。例如：

“旦”的理据是像太阳（日）从地平线（一）升起。“架”的理据是从“木”、“加”声。

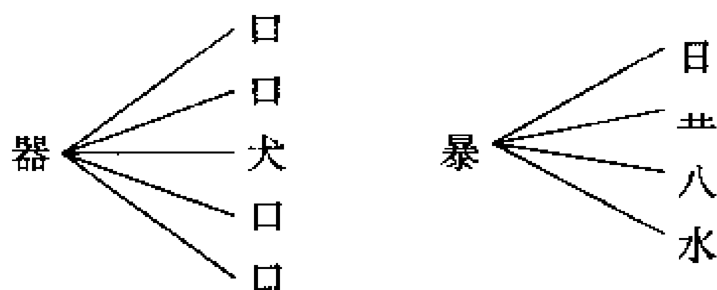
现代汉字中有一部分字无法直接分析造字意图，或形体与字源发生矛盾而无法参考字源分析理据。

（2）结构层次。汉字按理据由多个部件组合时，大部分是依层次进行的。在每个层次中，一般为两部件组合。少部分为一次性多部件平面组合。例如：

“蓓”、“蕾”为多层次组合，各层均为两部件。



“器”、“暴”为一次性多部件平面组合。



2. 部件拆分

将汉字拆分为部件称部件拆分。

（1）拆分依据。拆分依据是指将汉字拆分为部件的理论或现实根据。根据结构所进行的部件拆分，称有理据拆分；当无法分析理据或理据与字形发生矛盾时，依照字形所进行的部件拆分，称无理据拆分。

（2）拆分程序。拆分程序是指将汉字拆分为部件的次序。对层次结构的汉字递次进行有理据拆分，称层次拆分；对平面结构的汉字进行一次性的有理据拆分或无理据拆分，称平面拆分。

（3）过渡部件。过渡部件是指进行汉字层次拆分时，在完成最后一步拆分之前所拆分出的合成部件，称过渡部件。

3. 部件归纳

将部件归纳成组的工作称部件归纳。部件归纳分同形部件归纳、主形与变体的归纳和形似部件归纳三方面。

（1）同形部件。形体完全相同的部件，不论其意源是否相同，均称同形部件。

（2）主形部件。在各组中具有代表和称说其他部件作用的领先部件称主形部件。

（3）变体部件。与主形意源相同、形体略有差异的部件称变体部件。常见的变体部件有：

①因部位不同或组合方式不同而书写发生变异的变体部件。例如：

“爪”，在上方写作“爫”；“川”，在下方写作“巜”，在中间写作“巛”。

②在组字时省减某些部件或笔画而与主形功能仍然相同的变体部件。例如：

“衣”与“衤”；“豕”与“豸”。

③在组字时重加某些部件或笔画而与主形功能仍然相同的变体部件。例如：

“长”与“長”。

④放置方向不同而与主形功能相同的变体部件。例如：

“八”与“丷”。

（4）形似部件。与主形部件形体相似，而意、源不同的部件称形似部件。例如：

“子”与“孑”、“孓”。

（5）部件组。综合考虑部件之间的形、意、源关系，将主形、变体和形似部件归纳在一起所形成的组合称部件组。部件组是部件系统中最基本的组合。

（四）部件拆分的原则

1. 不违背形体的原则

（1）交重不拆。例如：

“串”不拆分为“中”、“中”；“东”不拆分为“七”、“小”。

（2）相离相接可拆。例如：

"明"拆分为"日"、"月"（相离）；"名"拆分为"夕"、"口"（相接）；"韭"拆分为"非"、"一"（相接）。

（3）极少数不影响结构和笔数的笔画搭挂，按相接处理。例如：

"孝"为"子"搭挂在"耂"上，可拆出。

2. 遵照结构规律的原则

（1）字形合乎理据的，进行有理据拆分。无法分析理据或形源矛盾的，依形进行无理据拆分。例如：

"分"拆分为"八"、"刀"（根据理据，有理据拆分）；"赤"拆分为"土"、"𣥂"（参考字源，有理据拆分）；"亦"拆分为"亠"、"𣥂"（无法分析理据，无理据拆分）；"虎"拆分为"虍"、"几"（形源矛盾，无理据拆分）。

（2）在进行多部件的有理据拆分时，应依汉字组合层次进行。例如：

"蓓"的第一次拆分为"艹"、"倍"，第二次拆分为"艹"、"亻"、"咅"第三次拆分为"艹"、"亻"、"立"、"口"。

3. 相关对应的原则

（1）有对应关系的简化部件与繁体部件一般应统一处理。例如：

"讠"与"言"（简体不拆分，繁体也不拆分）。

（2）主形与变体，尽量统一处理。例如：

变体"肉"与主形"月"均不拆分。

4. 合理控制下限的原则

（1）两个以上的相离、相接部分，若其中任一个无构其他字的能力，则不再拆分。例如：

"西"不拆为"一"、"西"（"西"无构其他字能力）；"黑"不拆分为"里"、"灬"（"里"无构其他字能力）。

（2）两个相离的对称部分，不分开用以构字的，不再拆分。

例如：

“兆”、“竹”、“非”不再拆分。

（3）传统独体字和部首，以及构字能力较强或拆分后均为非字部件的，虽有相离或相接部分，不再拆分。例如：

“石”不拆分为“丆”、“口”；“禾”不拆分为“丿”、“木”。

（4）拆分出的部件中，有一个是单笔画的，一般应为有理据拆分。例如：

“太”拆分为“大”、“丶”；“生”拆分为“丿”、“土”；“犬”不拆分为“大”、“丶”；“自”不拆分为“丿”、“目”。

（5）最末一次拆分，不可拆成两个以上的单笔画。例如：

“刀”不拆为“㇆”、“丿”；“刂”不拆为“丨”、“亅”；“川”不拆成“丿”、“丨”、“丨”；“兴”、“氵”均不再三分；“小”、“灬”均不再四分。

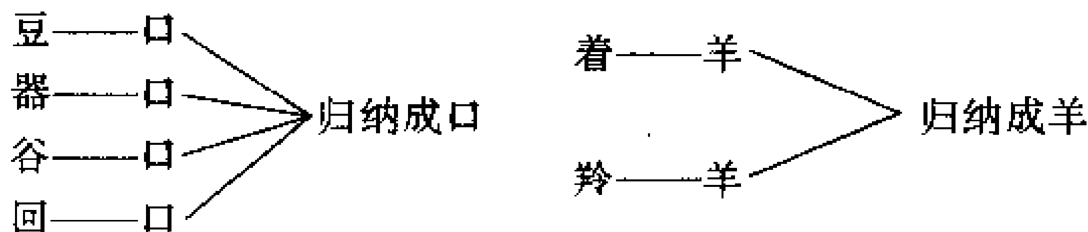
5. 综合衡量的规则

以上规则无法兼顾时，应权衡其利弊，综合处理。

（五）部件归纳的原则

1. 按形归纳的原则

（1）处在不同部位、大小不同的同形部件，均归纳为一个，不再分立。例如：



（2）笔画变异的同一部件，变异的笔画与原笔画同类者，归类为不分立的部件；笔画不同类者，按变体归纳。例如：

“火”作左面偏旁时，捺改写为点，不再分立；“王”作左面

偏旁时，横改写为提，不再分立；“木”中间一笔或为竖，或为竖钩，不再分立。

（3）意、源虽相同，但形体差异较大的部件，分立为不同部件。例如：

“火”与“灬”分立；“水”与“氵”分立；“手”与“扌”分立；“心”与“忄”分立。

2. 兼顾传统的原则

（1）传统独体字作部首时，因部位不同而写法相异，除形体差异较大者外，一般按主形与变体归纳。例如：

“示”（主形）与“礻”（变体）；“金”（主形）与“钅”（变体）；“食”（主形）与“饣”（变体）；“衣”（主形）与“衤”（变体）。

（2）将意源不同的同形部件归纳为同一部件时，用不同的例字提示其部位与来源的差别。例如：

“口”的例字，“吐”、“扣”、“司”、“占”中“口源于口舌之口；“吊”中之口为讹变之形；“高”、“兽”、“豆”中之“口”源于象形号。

”厶”的例字，“私”、“公”中的“厶”同源不同部位；“丢”、“允”、“勾”、“牟”中的“厶”均为不同位、不同源的同形部件。

3. 简化部件系统的原则

（1）有对应关系的简体与繁体部件，按主形与变体归纳。例如：

“贝”与“貝”。

（2）组字字数极少的罕用部件，尽量避免单独成组。例如：

“孑”、“孓”依形归入“子”组。

（3）日、韩方提出的汉字的部件，不论归在哪组，均以“＊”标识。

4. 优选主形的原则

（1）非字部件与成字部件之间，以成字部件为主形。例如：“八”（主形）与“丷”（变体）；“文”（主形）与“攵”（变体）。

（2）繁体部件与简体部件之间，以简体部件为主形。例如：“见”（主形）与“見”。

（3）上述两个条件无差别时，以组字能力较高、便于称谓的部件为主形。例如：

“土”（主形）与“士”（形似部件）；“子”（主形）与“孑”、“孓”（形似部件）；“人”（主形）与“入”（形似部件）。

（六）基础部件表的使用规则

《信息处理用GB13000．1字符集汉字部件规范》中的《汉字基础部件表》是对GB13000．1字符集中的20 902个汉字逐个进行拆分、归纳与统计后制定的。表中共提供基础部件393组，包括主形部件和附形部件合计560个。详见附件《汉字基础部件表》和《笔画排序汉字基础部件检索表》附表。

《汉字基础部件表》的使用规则有以下两点：

1. 基础部件不再拆分的规则

（1）《汉字基础部件表》中的部件均为基础部件，不得再行拆分。例如：

“非”不得再拆分为左、右两部分；“矛”不得再拆分为“マ”和“才”。

（2）《汉字基础部件表》中的部件没有包容关系，不得将大部件拆分后归入小部件。例如：

“疒”不是“广”与“冫”的组合，不得再拆分为“广”和“冫”；“广”不是“厂”与“丶”的组合，不得再拆分为“厂”和“丶”。

2. 基础部件可以组合的规则

《汉字基础部件表》中的基础部件可组合为成字部件使用，但不得组合出非字部件使用。例如：

可用“龷”、“由”、“八”组合成“黄”，作为部件使用；可用“自”、“田”、“丌”组合成“鼻”，作为部件使用；不得用“龷”、“由”组合成“寅”，作为部件使用；不得用“自”、“田”组合成“畠”，作为部件使用。

第四节　汉字信息处理的未来发展

本章就信息处理过程中涉及到的汉字输入问题进行了讨论，并相应介绍了国家语言文字工作委员会最新颁布实施的《信息处理用GB13000.1字符集汉字部件规范》标准。从更广义的学科角度上讲，汉字与信息处理的内容远不止这些。仅从人工智能与自然处理的研究与发展领域来看，就已经产生汉语文本的分析与生成、汉字识别、语言识别与合成等方面的研究课题，而且这些研究课题在某种程度上已经取得较大的进展，以汉字识别为例，对印刷体汉字和手写体汉字识别的研究取得了实质性的突破，而且有很多实用性很强的产品投入市场，如印刷体与手写体汉字识别系统、语音识别系统、中文文稿校对系统等。这些产品的成功使用，一方面刺激了市场需求；另一方面也促进了这一研究领域的发展，使这些领域的研究引起更广泛的关注，这对学术发展与技术进步都是非常有益的。

但另一方面我们也应看到，由于一段时期以来，在世界范围内人工智能和自然语言处理整个研究领域的工作都处于一个低潮，这或多或少也影响到了汉字信息处理，要使汉字信息处理达到人们所期望的理想程度仍有漫长的路要走。

随着多媒体计算机时代的到来，纯以文字个体作为信息载体来阐述某种系统思想、方法，已经不能满足时代的需求了，因此可以通过声音、动画、图象、影像等多种媒体来表征。中华民族拥有五千多年的文明史，这些文明通过浩如烟海的古代典籍以文字的形式传承下来，继承和发展这些文明是当代人不可推卸的责任。但是如何继承、如何发展则不是仅凭一腔热血就可以实现的，需要认真、扎实的学术研究。

用系统论的方法对古代典籍的系统思想进行研究，并用系统工程化的方法通过多媒体计算机对这些系统思想进行阐述和表征，这对于我们理解先人的思想、继承人类文明以及促进社会发展都是十分必要的，仅从这一点来看，汉字信息处理在这一研究领域就有着十分广阔的天地可供驰骋。

反过来也应看到，信息处理技术也正在促进语言文字的研究和发展。信息技术不仅将改变我们的思想观念和研究方法，而且也必将对整个世界的经济结构、社会结构乃至生活结构产生深刻的影响和变革。重视这种发展趋势，同时也重视汉语汉字本身的特异性质，才能在汉字与信息处理的研究与应用方面取得进展。

附表一　汉字基础部件表

序号	组号	部件	例　字	序号	组号	部件	例　字
1	1	口	吐扣司占吊高兽豆	41		(ｔ)	切
2	2	一	元旦与冃	42		(匕)	仑它尼匙北旨比
3	3	八	扒分兑兴夔	43		(上)	比顷
4		(丷)	半兑来曾	44		(七)	化
5	4	木	村林杏來果困	45		(ㅓ)	叱
6		(朩)	亲余寨	46	20	又	友双叉凤劝仅发难圣
7		(朮)*	朮瘸	47		(㇇)	祭
8	5	日	时昌旦早旬复者	48	21	亠	亩亥
9		(曰)	汨	49	22	大	太关夾夺
10		(冃)	冒冕	50	23	止	址正企步歧
11		(冃)*	冃胄	51		(止)	延
12	6	人	从因欠队谷合亥傘	52		(疋)	足疋走
13		(入)	籴歈	53		(㇀)*	疌
14	7	土	地圭吐圣去啬	54		(𠃊)*	懸
15		(士)	吉壮志增	55	24	扌	提
16	8	艹	草宽	56	25	女	妹妆囡妄汝
17	9	亻	仁亿	57	26	⺈	角争免
18	10	氵	河衍	58	27	宀	宝穴空
19	11	月	胖朋朗服	59	28	丶	主义凡叉乓刃勺
20		(月)	青有背	60	29	丿	乇么失乒
21		(月)*	臆盼	61	30	言	議信辩訇䛶
22		(夕)	炙祭將然	62		(讠)	论
23		(㐄)*	亙	63	31	冖	冠冪罕
24	12	灬	然庶盡	64	32	虫	虹萤浊
25	13	金	鍚淦鉴	65	33	火	灯炎伙灭灰灵
26		(钅)	银	66	34	冂	同网
27	14	十	什午协卖南	67		(冂)	周
28	15	田	畔畜画男佃思畏	68		(冖)	奂奥
29	16	纟	红	69	35	丷	前并豆
30		(糹)	紅	70	36	山	峡崇岔仙岁
31		(糸)	紫系絲辮	71	37	厶	私公丢允厷勾牟
32		(纟)	丝	72	38	王	珠主旺闰弄匷呈
33		(幺)	幼玄後幽	73		(玉)	国宝钰
34	17	文	纹刘吝斐	74	39	勹	包勺
35		(攵)	攻變	75	40	心	想沁闷慶
36		(夂)	冬麦	76		(⺗)	恭慕
37		(夊)	夊	77	41	小	孙少尖京尔
38	18	贝	财质则员狈	78		(⺌)	尚肖光
39		(貝)	購贊	79	42	阝	阳都
40	19	七	柒皂乇	80	43	立	站竖泣拉

序号	组号	部件	例字	序号	组号	部件	例字
81	44	禾	和秋秀颖秦乘	124		(刁)	叼
82	45	鸟	鸟岛枭	125	75	巳	包导氾
83		(鳥)	鳥梟鬼	126		(已)	已
84	46	衣	依袋裁	127		(己)	记忌岂
85		(衤)	补初	128		(㔾)	仓卷犯
86		(𧘇)	褰哀表	129		(己)	改凯
87	47	忄	快怪	130		(巳)	敃
88	48	圭	佳奮権	131		(巳)	顾
89	49	竹	竹	132	76	尸	层尾卢
90		(⺮)	简笑	133	77	爪	抓笊爬
91	50	目	睛督相盾瞿	134		(爫)	采妥受奚
92	51	辶	进连	135	78	彡	形杉彦须髟
93	52	几	机凡朵亢	136	79	子	李团享好孔
94	53	石	碗磨泵	137		(孑)	孑
95	54	寸	付尊村	138		(孓)	孓
96	55	囗	圈回	139	80	疒	病
97	56	车	轮军库轰	140	81	耳	聪取聋洱聂
98		(車)	輔輿	141	82	卜	仆赴
99	57	刂	刚别	142		(⺊)	占卢支卤
100		(刂)	斑辨	143	83	羊	样羣氧
101		(刂)	师	144		(⺶)	羚翔差着
102		(刂)	临坚	145		(⺷)	美羔義
103	58	厂	雁厦岸	146	84	冫	冰匀
104	59	四	泗	147		(冫)	冬寒尽头斗於
105		(罒)	罗曼	148	85	兀	元光尧匾
106	60	𠂉	午每矢	149		(尢)	尤无尬
107	61	广	廣圹庄	150	86	儿	兒匹兢
108	62	白	伯帛皤百	151	87	犭	犹狗
109	63	工	江巧式贡巫	152	88	丰	半击用奉
110	64	乂	艾交义风区爽	153		(丰)	判
111	65	马	马	154	89	牛	件牢牲
112		(馬)	馬凴	155		(⺧)	告先
113	66	力	肋助办劣历边	156	90	凵	凶出击画
114	67	皿	盘血	157	91	习	羽
115	68	𠂇	右发戎	158	92	戈	伐戆划戛
116	69	米	料粟迷娄类	159	93	亏	号考亏兮巧
117	70	示	宗祘标票	160		(丁)	可
118		(礻)	祖视	161	94	龷	共昔巷暴
119	71	门	闪问们	162	95	雨	雪扁
120		(門)	閑閎	163	96	廾	弄开升
121	72	丆	页面	164		(廾)	廾
122	73	巾	布帆吊匝帅	165	97	𠂊	欠亇
123	74	刀	刃分切召	166	98	一	买疋

序号	组号	部件	例　字
167		（⺈）	今疏
168	99	彳	很街
169	100	夕	名多外宛奖罗舛
170	101	夫	扶规
171	102	方	旅仿房
172	103	食	飧餐
173		（飠）	飽
174		（饣）	饭
175	104	酉	酊酒酋
176	105	水	冰凼氽沓
177		（氺）	泰暴录黍
178		（氺）	犀裏屬
179		（氺）	眔
180	106	龶	青素毒责
181	107	虍	虎虏虛
182	108	乚	孔亂甒
183		（乚）	甚断匃陋眞県値
184	109	豕	家圂豬逐豚
185		（豸）	象彖
186		（豕）	啄冢剢瘃
187	110	二	式些仁竺
188	111	川	训顺
189		（⺦）	流侃荒
190		（巛）	巡巠
191		（巛）*	淄
192		（巜）	粼俞
193	112	干	旱刊肝
194		（干）	幵
195	113	匚	区叵汇
196		（匚）	匚
197	114	聿	律筆晝
198		（聿）	書畫
199		（肀）	肅
200		（肀）	唐肃
201		（聿）	盡賷
202		（肃）	肅
203		（肃）	肃
204		（肀）	隶
205	115	西	栖茜迺
206		（覀）	要
207		（襾）	襾
208	116	斤	新匠斧颀
209	117	臼	兒舀舌
210	118	艮	很垦良
211		（艮）	即既郎
212	119	彐	雪归当寻
213		（彐）	录
214		（彐）*	婕
215	120	丰	争艳拜書
216		（丰）	寿邦
217		（丰）*	丰
218	121	见	视规觉
219		（見）	現覓
220	122	弓	引躬弩
221	123	戊	戌成威
222		（戊）	越钺
223	124	耂	老考者
224	125	丁	钉亭厅
225	126	犬	伏器狱
226	127	屮	出蚩艸
227		（屮）	逆舞艸辥
228	128	牛	韋舛降
229	129	户	启妒芦
230		（戶）*	戶
231		（戶）	所戶
232	130	革	鞋霸缂
233	131	𠂉	丧畏展
234		（𠂢）	旅派
235	132	丨	旧候
236		（亅）	寧亇
237	133	由	油宙迪專黄
238		（甴）	甴
239	134	长	帐套
240		（長）	張
241		（镸）	肆髟套
242	135	𠃌	勿昜
243	136	⺈	令甬亂疑
244	137	舟	船俞盘
245	138	𢉖	鹿慶麃
246	139	卩	节印卫卵卿
247		（㔾）	报
248	140	歹	列夙
249	141	册	删姗箭
250		（冊）*	刪姍珊
251		（冊）	扁侖嗣龠
252	142	冎	骨

序号	组号	部件	例　字	序号	组号	部件	例　字
253		(冎)	咼	296	176	丬	壮将
254	143	𠂆	反后卮	297		(爿)	牀
255	144	而	耐耍需侕	298	177	單	单單
256	145	甘	其	299	178	世	屜抴枼
257	146	臣	卧姬宦臧	300	179	皮	坡疲颇
258		(𦣝)*	鑒	301	180	黑	嘿墨默
259	147	业	業显凿	302	181	亡	忙忘氓望罔
260		(业)*	虛	303	182	瓦	佤瓩瓷
261	148	未	味朱	304	183	⺍	兴学举金
262		(末)	昧茉	305	184	㐄	韋
263	149	自	鼻咱息臬首	306		(龶)*	麦
264	150	非	匪韭排剕罪	307	185	𡗗	春奉秦
265	151	毛	笔尾毡毫蚝	308	186	也	他𥬲
266	152	几	凤	309	187	丸	执芄
267	153	巴	吧爸邑	310	188	不	丕吥罘还
268	154	廿	堇度	311	189	内	纳丙氖
269	155	龷	寒冓襄	312		(内)*	呐
270	156	且	助宜咀县	313	190	父	斧爸
271	157	矛	茅務柔	314	191	⺫	曾會
272	158	缶	缸陶窑	315	192	果	巢夥棵裹
273	159	千	仟臿乖舌	316	193	上	让叔
274	160	中	仲盅衷串	317		(丄)*	丄
275	161	甫	俌圃庸尃莆	318	194	少	步賓歲
276	162	氏	纸昏氐	319	195	𠂎	留劉
277	163	川	齐氚	320		(丣)	柳
278		(川)	肃渊	321		(丣)*	卯
279	164	𠫔	至	322		(丣)*	酉
280	165	束	辣速	323	196	釆	番释悉
281		(朿)*	枣	324	197	鬼	塊瘣魅
282	166	里	哩厘童野	325	198	𦥑	學舉輿
283	167	廴	延建	326	199	亚	哑恶壶氩
284	168	母	姆每毓繁纛毐	327		(亞)	啞惡氬
285		(毋)	毒	328		(亜)*	唖悪
286		(毌)	贯	329	200	身	射躬
287	169	弋	式貳	330		(㐆)	殷
288	170	手	掌拿	331	201	乃	仍疠氖秀
289		(手)	拜掰看	332	202	免	免兔谗
290	171	氺	隶函脊率	333		(兎)*	兎
291	172	龙	拢垄宠	334	203	𠔉	兼廉
292		(𦚏)	龍	335	204	柬	練
293	173	癶	登癸發	336		(东)	拣练
294	174	乙	亿乞	337	205	禺	偶寓愚颙
295	175	且	具直縣	338	206	𠃌	司局幻

序号	组号	部件	例　字	序号	组号	部件	例　字
339	207	九	仇宄杂虓尯	382	244	弔	弟帏蛊
340	208	小	亦赤	383	245	五	伍吾
341		(小)	变	384	246	开	异亓
342	209	𠂉	临	385	247	冃	齊
343	210	电	黾奄	386	248	尸	眉声
344		(电)	黿奄	387	249	目	官粗首
345		(黽)*	龟	388	250	民	[illegible]
346	211	气	氣氧忾	389	251	㇏	乆
347	212	彑	彝录彖彙叠	390	252	乛	敢島
348	213	甶	卑	391	253	丫	齊吖
349	214	万	厉迈	392	254	内	离
350	215	与	与焉鳶	393	255	申	伸畅审
351		(与)*	舄	394	256	⺕	争
352	216	冉	再冓	395	257	㡀	敝蔽鼈
353		(冄)*	髯	396	258	于	玗宇盂迂
354	217	央	映盎英	397	259	下	吓苄
355	218	[illegible]	鼠	398		(下)*	丅
356	219	豸	豹	399	260	𠂆	齊
357	220	曲	軸農	400	261	冘	沈鸩
358	221	𠂤	阜追師薛	401	262	东	冻陈
359	222	旡	既潛炁[illegible]	402		(東)	棟東鶇氭
360		(旡)	旡	403	263	[illegible]	假
361	223	耒	耕诔	404	264	[illegible]	高回
362	224	兆	佻晁逃	405	265	夬	决
363	225	乍	作岝厏	406	266	[illegible]	堇
364	226	必	泌苾邲	407	267	乡	鄉雍
365	227	甘	某甜	408	268	及	汲岌
366	228	叀	惠專	409	269	更	便郠甦
367	229	夕	亥夂遠	410	270	[illegible]	淵
368	230	㒳	兩爾两	411	271	户	倉
369	231	屯	纯囤笔顿	412	272	予	序妤预
370	232	[illegible]	禹	413	273	甲	匣呷閘鸭
371	233	[illegible]	[illegible]	414		(甲)	早
372	234	巨	距苣	415	274	[illegible]	叟
373	235	牙	呀穿邪	416	275	片	版拚
374	236	瓜	弧瓢瓞瓣	417	276	为	伪
375	237	了	亨疗钉[illegible]	418		(爲)	僞爲蔿鄬
376	238	コ	假叵兜	419		(為)	[illegible]蔿
377	239	弗	佛氟第艳	420	277	丑	妞羞
378	240	我	俄義鵝	421	278	卌	無
379	241	重	動僮董衝	422	279	[illegible]	鹵[illegible]傻
380	242	尹	伊君笋	423	280	求	救俅裘逑
381	243	丘	邱乒蚯	424	281	垂	睡甀菙

序号	组号	部件	例字
425	282	黽	繩鼉
426	283	𠂎	卯卿
427	284	𠁣	曹
428		(亩)*	曹叀專惠
429	285	尺	迟尽呎咫
430	286	𠌶	華
431	287	乁	虱
432	288	𡿨	匆囱
433	289	韦	伟围违韬
434	290	卅	卅
435		(卅)	带
436		(卌)	帶
437		(卅)*	帶
438	291	乎	呼虖
439	292	戈	尧
440	293	丹	坍彤旃
441		(円)*	青
442	294	𠃓	杨疡飏
443	295	戋	钱浅盏
444		(戔)*	錢賤
445		(𢦏)	彧
446	296	乌	乌
447		(烏)	烏
448	297	朿	刺棘枣
449	298	曲	典
450	299	夷	嫫冀瘼鷞
451	300	亠	侯
452	301	𢀖	茎经劲
453	302	疌	捷婕
454	303	畢	嗶簞
455	304	久	玖灸疚羑
456	305	𠂉	衣雇
457	306	之	芝乏
458	307	井	耕进
459	308	莫	漢
460	309	勿	黎犁
461	310	禹	属瑀萬
462	311	熏	勳薰獯
463	312	丩	收叫
464	313	本	体笨梆
465	314	𠀉	囊橐叓
466	315	肉	腐
467	316	串	患窜腓
468		(弗)	弗
469	317	永	泳昶羕
470	318	亘	覆夓
471	319	巨	虐疟
472		(㔾)	印寝
473		(巳)*	仓
474	320	市	芾沛
475		(朩)*	柿
476	321	𠂔	妻
477	322	𠂤	段
478	323	臣	颐宧
479		(臣)*	熙
480	324	臾	谀庾
481		(臾)*	萸
482		(臾)*	䫻
483	325	歺	餐
484	326	龙	襱
485	327	象	像勨
486	328	才	財豺团闭
487	329	冂 *	禹
488	330	两	俩魉
489	331	𠂆	卬兜
490	332	𠄎	丞
491	333	缶	御衔
492	334	門	閑
493	335	𠃊	以
494	336	斥	坼
495	337	肀	庚
496	338	𠂉	制
497		(𠂉)*	㡀
498	339	州	洲酬
499	340	𡿬	鼎
500	341	𠀃	襄
501		(𠀃)*	囊
502	342	𠂔	姊
503	343	乐	烁
504	344	曳	拽曳
505	345	乑	衆聚
506	346	戢	濈戢
507	347	三	仨叁式闫
508	348	非	非啡韭
509	349	夂	夜
510	350	史 *	剑俭

序号	组号	部件	例字
511	351	ㄣ *	吴呉
512	352	𠂇	发
513	353	丒 *	㑥
514	354	飞	飞
515		(飛)	飛
516	355	专	传
517	356	丈	仗
518	357	丏	沔[illegible]
519	358	冄	那
520	359	年	哖[illegible]
521	360	[illegible]	[illegible]
522	361	乀	[illegible]
523		(乀) *	心
524	362	旡 *	[illegible]
525	363	丐	钙
526	364	史	駛
527	365	丱	聯關
528	366	[illegible]	[illegible]
529	367	事	倳剚
530		(事) *	[illegible]
531	368	[illegible]	[illegible]
532	369	力 *	別
533	370	夨	奊[illegible]
534	371	凹	兕[illegible]
535	372	[illegible] *	[illegible]
536		([illegible]) *	[illegible]
537	373	亐 *	污
538	374	乇	甩
539	375	冂	滿
540	376	罒	临
541	377	卍	卍
542		(卐)	卐
543	378	吏	使
544	379	隶	棣
545	380	乘 *	剩
546	381	甫 *	捕
547		([illegible]) *	[illegible]
548	382	亜	壺
549		([illegible]) *	[illegible]
550	383	乂 *	乂
551	384	乜	乜
552	385	个	个
553	386	书	书
554	387	[illegible] *	[illegible]
555	388	凸	凸
556	389	壴 *	[illegible]
557	390	業	業
558	391	承	承
559	392	西	[illegible]
560	393	乇 *	托

附表二　笔画排序汉字基础部件检索表

笔数	首笔	部件	序号	笔数	首笔	部件	序号	笔数	首笔	部件	序号	笔数	首笔	部件	序号
1	一	一	2	2	丨	冂	68	2	㇇	丩	463	3	一	大	49
	丨	丨	235		丿	𠂉	106			ㄐ	111			丈	517
		亅	236			亻	17			力	113			兀	148
	丿	丿	60			丿丨	277			刀	123			尢	149
	丶	丶	59			丿丨	278			乃	331			丘	471
		㇏	389			𠃌	242			ク	367			巨 *	473
		㇂	522			厂	254			ケ	512			牛	228
		乀 *	523			八	3			又	46			万	349
	㇇	㇖	166			人	12			ヌ	47			𠂇	439
		㇇	167			入	13			厶	71			云	279
		乙	294			乂	110			マ	243			弋	287
		𠃌	338			匕	42			廴	283			𠂉	533
		㇕	390			𠤎	44			ス	452		丨	⺕	45
		㇈	431			𠂊	57			乄 *	550			上	316
		㇄	182			⺆	67			卩	493			止	51
		㇗	183			丿亅	74			幺	32			少	318
		㇉ *	511			几	93			巳	128			小	77
2	一	二	187			儿	150			㔾	131			⺌	78
		十	27			𠂉	165			巜	192			口	1
		丁	160			⺇	266			乜	551			囗	96
		丁	224			九	339	3	一	三	507			山	70
		丅 *	398			亐 *	351			丰	152			巾	122
		厂	103			匚	489			干	193			冂	539
		𠂇	115			力 *	532			于	396		丿	千	273
		丆	121		丶	亠	48			丯	153			𠂉	342
		七	40			冫	146			亍	194			乇 *	560
		七	41			丷	4			亐 *	537			川	188
		⺊	43			⺉	100			乇	538			巛	189
		丂	159			⺀	147			土	14			彳	168
		匚	195			讠	62			士	15			彡	135
		匸	196			冖	63			扌	55			夂	432
		与	350		㇇	刁	124			工	109			个	552
	丨	⺊	142			コ	376			艹	16			巨	472
		丄 *	317			ユ	451			廾	164			犭	151
		刂	99			阝	79			才	486			夕	426
		刂	102			凵	156			寸	95			勿	460
		刂	101			卩	246			下	397			夂	36
		卜	141			⺋	247			廾	163			夊	37
		冂	66			了	375			丌	384			夕	169

笔数	首笔	部件	序号	笔数	首笔	部件	序号	笔数	首笔	部件	序号	笔数	首笔	部件	序号
3	丿	丸	309	3	㇇	飞	514	4	丨	止	50	4	丿	攵	86
		𠂆	399			习	157			龰	52			⺀	234
		及	408			㣺 *	54			冃	404			月 *	21
		久	455			纟	29			⺗	341			丹	440
		饣	174			⺓	347			小	76			夕	22
		乌	446			乡	407			日	8			夕 *	23
	丶	丬	296			幺	33			曰	9			夂	509
		广	107			巛	190			⺜	10			鸟	82
		亡	302	4	一	丰	215			冃 *	11		丶	文	34
		门	119			丯	216			月	20			⺌	290
		䒑	69			王	72			𦘒 *	353			方	171
		丫	391			⺺	180			卄	500			䒑	370
		⺗ *	53			井	458			中	274			火	65
		氵	18			朩	6			冉 *	441			为	417
		忄	87			夫	170			罒	540			灬	24
		⺍	304			无 *	524			冂 *	487			户	229
		宀	58			韦	433			贝	38			礻	118
		亠	456			专	516			水	176			尤	400
		辶	92			耂	223			内	311			心	75
		之	457			丐	525			内 *	312		㇇	⺺	200
	㇇	彐	212			丗	161			见	218			⺻	204
		⺕	213			廿	268			內	392			⺕	394
		⺔ *	214			卝	508		丿	牛	154			尹	380
		肀 *	306			卌	435			手	288			戸	411
		尸	132			木	5			龵	289			⺳	519
		巳	125			𣎳 *	7			毛	265			夬	405
		已	126			五	383			气	346			尺	429
		己	127			而	368			⺧	155			弔	382
		㔾	129			市	474			攵	35			丑	420
		⺋	130			丏	518			长	239			⺸	542
		弓	220			卅	434			⺼	385			卍	541
		子	136			不	310			片	416			爿	297
		孑	137			犬	225			⺌	340			尸	386
		⺪	305			歹	248			斤	208			巴	267
		屮	226			戸 *	230			爪	133			予	412
		屮	227			巨	372			⺷ *	217			毌	286
		孓	138			车	97			戶	231			毋	285
		也	308			牙	373			父	313			书	553
		女	56			屯	369			爫	134			𠄎 *	191
		刍	442			戈	158			乑 *	554	5	一	⺶	307
		𠃜	233			旡	359			月	19			玉	73
		癶	490			旡 *	360			氏	276			龷	269
		丮 *	513			瓦	303			⺸	502			廿 *	501

笔数	首笔	部件	序号	笔数	首笔	部件	序号	笔数	首笔	部件	序号	笔数	首笔	部件	序号
5	一	示	117	5	丨	罒	105	6	一	臣	257	6	丶	羊	145
		未	261			皿	114			吏	543			米	116
		末	262			冊 *	250			覀	206			州	498
		市 *	497			凹	534			襾	207		㇇	聿	197
		戋	443			四	104			束	448			聿	198
		甘	365			冎	253			西	205			肃	203
		世	299			罒	252			西	559			良	210
		丗 *	437	丿		钅	26			丗	436			糸	31
		卌	421			乍	363			而	255			糹	30
		本	464			禾	81			夷	450	7	一	垂	430
		朩 *	475			阝	477			戋	445			镸	241
		毌	465			丘	381		丨	卢	181			亘 *	258
		石	94			白	108			业 *	260			丽 *	322
		戊	221			斥	494			且	295			臣	478
		龙	291			瓜	374			早	298			臣 *	479
		肀	476			乎	438			吏 *	481			求	423
		东	401			乐	503			臾 *	482			車	98
		戉	222			册	249			曳	504			甫	275
		东	336			⺳	319			虫	64			更	409
	丨	凸	555		丶	疒	139			曲	357			亜 *	328
		歺	483			立	80			曲	449			曲	427
		业	259			礻	85			罒	314			束	280
		氺	179			必	364			肉	466			朿 *	281
		氺	177			永	469		丿	缶	272			两	488
		氺	178		㇇	聿	199			㐬	496			酉	175
		目	91			聿	201			年	520			卵	320
		且	270			聿	495			竹	89			豕	184
		申	393			艮	211			⺮	90			虍	484
		甲	413			阝	403			氶	505			事 *	530
		甲	414			民	388			臼	209			亩	292
		电	343			弗	377			囟	263		丨	里	282
		电	344			丱	527			甶	348			虫	406
		田	28			皮	300			鸟	447			貝	39
		由	237			癶	293			自	358			見	219
		冉	352			矛	271			⺼	330			串	467
		冊	251			母	284			舟	244		丿	缶	491
		央	354	6	一	来	361			兆	362			我	378
		史 *	510			戋 *	444			米	422			垂 *	547
		史	526			丐	112			豕	185			𦥑	325
		兄	332			耳	140		丶	圭	88			由	415
		兄 *	333			其	256			衣	84			身	329
		目	387			亚	326			羊	143			鸟	83
		甴	238			亩 *	428			⺷	144			釆	323

笔数	首笔	部件	序号	笔数	首笔	部件	序号	笔数	首笔	部件	序号	笔数	首笔	部件	序号
7	丿	豸	356	9	丿	食	172								
	丶	言	61		丶	亞	548								
		产	245			為	419								
		亥	528		㇇	飛	515								
		㒳	395	10	一	鬥	492								
	㇇	隶	334		丨	畢	454								
		爿	499		丿	戴	531								
		丣 *	321		㇇	曹 *	535								
		凬	355	11	一	黄	459								
8	一	畫 *	556		丿	曾 *	536								
		長	240			象	485								
		亞	327	12	丨	戢	506								
		業	557			黑	301								
		叀	366		丿	爲	418								
		東	402	13	丨	黽	425								
		事	529	14	丿	熏	462								
		雨	162	16	丿	龜	521								
		冓	470												
		豖	186												
		疌	453												
	丨	非	264												
		果	315												
		害	371												
		典	468												
	丿	垂	424												
		秉	544												
		川	410												
		臾	480												
		金	25												
		飠	173												
	丶	亟 *	549												
	㇇	門	120												
		書	202												
		承	558												
9	一	革	232												
		柬	335												
	丨	黽 *	345												
		禺	337												
	丿	乘 *	545												
		重	379												
		垂 *	546												
		鬼	324												
		禹	461												